A PHYLOGENETIC PERSPECTIVE ON THE DYNAMICS OF SPECIATION AND EXTINCTION DURING EVOLUTIONARY RADIATIONS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Daniel Lee Rabosky

May 2009

A PHYLOGENETIC PERSPECTIVE ON THE DYNAMICS OF SPECIATION AND EXTINCTION DURING EVOLUTIONARY RADIATIONS

Daniel Lee Rabosky, Ph. D.

Cornell University 2009

One of the most striking features of the natural world is that some groups of organisms are stunningly diverse while many other groups are species-poor. Both intrinsic, lineage-specific traits as well as extrinsic ecological factors can influence species richness through their influence on speciation and extinction rates. Here I investigate intrinsic and extrinsic influences on the diversification process, using a joint theoretical and empirical approach. I focused on reconstructing patterns of species diversification from time-calibrated molecular phylogenetic trees, which provide a valuable window into macroevolutionary tempo and mode for the large number of groups with inadequate fossil records.

I review the manner in which species traits influence the dynamics of speciation and extinction, and I relate the literature on "species selection" to the emerging perspective on diversification rates from modern molecular phylogenetic studies. I argue that this literature demonstrates the effects of species selection on both species richness and the distributions of phenotypic traits in higher taxa. I develop tools for analyzing variation in speciation and extinction both over time and among lineages, including a new analytical approach for modeling rates that vary continuously through time. Using several real phylogenies and simulations, I show that, under the birth-death model, the phenomenon of early rapid diversification in phylogenies can only be explained by decreasing speciation through time in

conjunction with low background extinction. I develop an alternative model that postulates constant diversity and a balanced speciation-extinction process that can also explain the rapid accumulation of lineages during the early stages of many radiations.

I found evidence for an explosive increase in diversification in a lineage of Australian scincid lizards, suggesting that the evolution of traits associated with climate tolerance may have played a role in shaping patterns of diversity in this group.

Using multiple published datasets, I tested three hypotheses for the absence of an age-diversity relationship in higher taxa, including (1) among-clade rate heterogeneity, (2) clade volatility, and (3) ecological limits on clade growth. Using a new modeling framework, I found that only ecological limits are capable of explaining the observed patterns.

BIOGRAPHICAL SKETCH

Daniel Lee Rabosky opened his eyes to the world on February 24, 1977, in a small town on the Ohio-Pennsylvania border. After several violent incidents involving his infant brother and his parent's dog Yeager, the young Dan-in-diapers seemed bound for a career in street thuggery or petty theft. However, he soon discovered that natural history was his true love. His earliest memory involved finding a crayfish claw in the Lake Erie surf at Presque Isle, Pennsylvania, when he was too small to stand up to the force of even small waves.

Perhaps his next memory is of the time a neighbor gave him a small common garter snake (*Thamnophis sirtalis*), which he named Blue Streak. Snakes promptly grew into an obsession that has not left him to this day. He spent his earliest days looking for fossils, snakes, salamanders, and other creatures near his home in rural Ohio. His parents, Lynn and Ron, actively encouraged him in what all of the neighbors surely thought was a most peculiar of pursuits. Neither of his parents were particularly keen on snakes, but they were more than tolerant: among other things, his father built him snake cages, and his mother would drive him to swamps and wait patiently while he slogged around in the muck for hours, only to return with a bag of *Nerodia* water snakes. The escape of a milk snake in the house forced him to be more seasonal in his herpetological ambitions, because limbless tetrapods were henceforth banned from his parents' house, and he could only maintain his collection outside during the warm summer months. Nonetheless, he managed during several summers to open a small roadside zoo; despite a bargain entrance fee of (in 1985 dollars) $0.10, the crowds did not follow and missed a unique opportunity to see the pet raccoons Winky and Bandit as well as a host of miscellaneous turtles, snakes, and other native beasts. Dan felt that a roadside zoo was a far better proposition than the lemonade stands of his schoolmates, but originality can be a tough sell in rural Ohio.

Though he did not appreciate it at the time, his value system was defined by those early days in Ohio. His dad, laid off from the railroad during the recession years of the early/mid ‘80s, taught him the value of good work and the necessity of rising to whatever challenges life throws at you. Dan distinctly remembers the years where the family survived on the shoulders of his father’s extreme capacity for hard work. His mother encouraged him to pursue his intellectual dreams, wherever they might lead: she taught him to read and spent an awful lot of time taking him to libraries and museums and rock shops. Moreover, she felt so strongly about education that she co-founded a school in large part to give her children an excellent education.

As Dan grew older, his passions expanded to hunting and fishing. Together with his brother Darren, his father Ron took them throughout northeastern Ohio and west/central Pennsylvania in search of smallmouth and largemouth bass, whitetail deer, wild turkey, ruffed grouse, walleye, and many other animals. The most eagerly anticipated days of his year were the three days he was allowed to skip school to go turkey and deer hunting.

One of the most important moments in his life came when he was a sophomore in high school, reading E. O. Wilson’s *The Diversity of Life*. There, amid the dazzling descriptions of tropical biodiversity and convergent evolution of continental faunas, Dan suddenly realized that “Evolutionary Biologist” was a real career, that people actually *were paid to do this*, and that this was in fact what he wanted to do with his life. Here was the unifying conceptual theme that could link all of his interests in fossils and faunas.

He finished high school, saw his first wild grizzly, and left for college at Ohio University. He studied the evolution of beetle digestive enzymes, desert lizard behavior, and the population genetics of shrews and mice as undergraduate research projects. After graduating, he had a brief stint at the University of Chicago where he

realized that *Drosophila* population genetics was not really what he wanted to do with his life, and he moved to the west coast to work as a field biologist. He lived on San Clemente Island for a year and, in the name of conservation, shot cats and poisoned rats (among other things). He worked briefly as an environmental consultant in southern California, and then moved on to a fabulous job as a herpetologist for the Utah Division of Wildlife. Every day involved tramping through spectacular canyon country of the extreme northeastern Mojave Desert, looking for tortoises and rattlesnakes and horned lizards.

Dan completed an MS at Penn State on Australian blindsnake systematics, which helped pave the way for his work on *Ctenotus*. After a few celebratory months traveling through Ecuador, he moved to Cornell in 2003 to start PhD work under the tutelage of Dr. Amy McCune and Dr. Irby Lovette. This was surely one of the best things that happened to him in life. He was able to develop his interests in evolutionary theory, macroevolution and biodiversity, and combine them with fieldwork in a remote and wondrous part of the planet. His advisors encouraged him to cast a broad intellectual net, to follow his passions, and to try and release himself from those activities for which he had little aptitude, particularly labwork. Dan will move to the University of California at Berkeley in July 2009, where he will begin a three-year Miller postdoctoral fellowship.

To Mom and Dad

ACKNOWLEDGMENTS

I am grateful for the support of numerous friends and colleagues during the past six years. My PhD advisors, Amy McCune and Irby Lovette, have been true mentors and I cannot thank them enough for everything they have done for me during my time at Cornell. In terms of my intellectual development, it is hard to imagine a more perfectly complementary pair of advisors. They created an ideal environment in which I was able to simultaneously pursue my passions for evolution, biodiversity, squamates, and wild desert landscapes. I also thank them for allowing – or rather, actively encouraging – me to follow a series of intellectual threads that often may have seemed loosely connected at best. I thank them for their inspiration and generosity – and for taking a gamble on me in the first place.

I also thank my committee members, Anurag Agrawal and Harry Greene, for much valuable feedback and discussion over the years. Among other items, I have greatly enjoyed co-organizing several seminars with Anurag, and I am honored to have had the incredible opportunity to teach herpetology with Harry.

I thank the many members of Irby's lab over the years for valuable interaction and assistance in the lab. Given my general awkwardness with pipettors and associated lab tools, the *Ctenotus* project might have been hopeless without Amanda Talaba's incredibly competent hand. I have benefitted much from both her skills in the lab as well as her encyclopedic knowledge of herpetological diversity. I also thank Laura Stenzler for keeping the lab free of chaos and open-toed sandals, and Suzanne Kates for patiently working through various disorganized piles of receipts on my behalf. I also thank the members of the Big Kids group for valuable discussion over the years. I thank the Lovette and McCune lab members with whom I've had the pleasure of interacting for the past few years, especially Becky Cramer, Mari Kimura, Vale Ferretti, Andrea Townsend, Katie Wagner, Rachel Vallender, Aurelie Coulon, and

Mandy Cass.

My work in Australia would not have been possible without the assistance of Mark Hutchinson, Steve Donnellan, Mark Cowan, and Julian Reid. I have learned much from all of these individuals. I also thank Paul Doughty, Brad Maryan, and Claire Stevenson at the WAM; Carolyn Kovach at the SAM; Steven Argus; Danny Stefoni (WA DEC); Peter Kendrick; Warren Mitchell; Sandy Crameri; and many other individuals who have contributed in some way to the logistics of working in the field on the other side of the world.

I also thank Teresa Griffin for assistance with grants; and Wesley Hochachka, David Winkler, Daniel Fink, and Steve Ellner for assistance with miscellaneous quantitative challenges. I also thank my fellow EEB grads Jamie Walters and Katie Wagner, whom I have used repeatedly as a sounding board throughout my time at Cornell. My former committee members Jim Marden and Steve Schaeffer (Penn State) kept me on a track that led to Cornell, and Hiroshi Akashi taught me computer programming, one of the most important skills I have learned.

I thank my future wife, Dr. Dominic Dawson, for her love and encouragement, and for her willingness to share the next phase of this grand adventure with me.

Finally, I thank my parents for everything they have done for me. This would not have been possible without their support.

TABLE OF CONTENTS

BIOGRAPHICAL SKETCH iii

DEDICATION vi

ACKNOWLEDGEMENTS vii

LIST OF FIGURES xi

LIST OF TABLES xiv

CHAPTER 1. REINVENTING SPECIES SELECTION WITH MOLECULAR PHYLOGENIES 1

CHAPTER 2. LIKELIHOOD METHODS FOR INFERRING TEMPORAL SHIFTS IN DIVERSIFICATION RATES 16

CHAPTER 3. LASER: A MAXIMUM LIKELIHOOD TOOLKIT FOR DETECTING TEMPORAL SHIFTS IN DIVERSIFICATION RATES FROM MOLECULAR PHYLOGENIES 50

CHAPTER 4. EXPLOSIVE EVOLUTIONARY RADIATIONS: DECREASING SPECIATION OR INCREASING EXTINCTION THROUGH TIME? 59

CHAPTER 5. HERITABILITY OF EXTINCTION RATES LINKS DIVERSIFICATION PATTERNS IN MOLECULAR PHYLOGENIES AND FOSSILS 86

CHAPTER 6. EXCEPTIONAL AMONG-LINEAGE VARIATION IN DIVERSIFICATION RATES DURING THE RADIATION OF AUSTRALIA'S LARGEST VERTEBRATE CLADE 116

CHAPTER 7. DENSITY DEPENDENT DIVERSIFICATION IN NORTH AMERICAN WOOD-WARBLERS 137

CHAPTER 8. ECOLOGICAL LIMITS ON CLADE DIVERSIFICATION IN HIGHER TAXA 161

APPENDIX I. SUPPLEMENT TO CHAPTER 6 191

APPENDIX II. SUPPLEMENT TO CHAPTER 7 198

APPENDIX III. SUPPLEMENT TO CHAPTER 8 204

REFERENCES 208

LIST OF FIGURES

Figure 1.1 Organismic selection and species selection on a phenotypic trait 4

Figure 1.2 Repeated effects of a trait on diversification 10

Figure 2.1 Phylogeny with five species illustrating parameters discussed in text 22

Figure 2.2 Type I error rates and ΔAIC_{RC} rejection criteria for rate-constant phylogenies under different extinction fractions 32

Figure 2.3 Relationship between ΔAIC_{RC} and the extinction fraction *a* when survival analysis is used to compute likelihoods 33

Figure 2.4 Power of birth-death likelihood and the γ-statistic to detect shifts in diversification rates over time 34

Figure 2.5 Power of birth-death likelihood and the γ-statistic to detect temporal decreases in diversification 35

Figure 2.6 Power of birth-death likelihood and the γ-statistic when diversification rates have decreased gradually under a model of density-dependent cladogenesis 36

Figure 2.7 Relative error in parameters estimated under the best fit rate-variable model for phylogenies simulated under four models of diversification 37

Figure 2.8 Analysis of diversification in Australian agamid lizards 38

Figure 3.1 Distribution of the ΔAIC_{RC} test statistic for 5000 rate-constant phylogenies of the same size as the *Enallagma* phylogeny 56

Figure 4.1 Phylogenetic tree illustrating parameters described in text 65

Figure 4.2 Models used to simulate phylogenies undergoing five-fold decline in net

diversification rate 74

Figure 4.3 Maximum likelihood estimates of speciation rates and extinction rates under the BOTHVAR model for three phylogenies discussed in text..... 75

Figure 4.4 Mean log-lineage through time curves for phylogenies simulated under alternative diversification scenarios..... 76

Figure 4.5 Distribution of γ statistic for phylogenies simulated with identical net diversification rates but with different speciation and extinction rates 77

Figure 5.1 High extinction creates the impression of accelerating diversification 88

Figure 5.2 Parameters in heritable extinction with pulsed turnover (HEPT) model ... 92

Figure 5.3 Posterior distributions of parameters in the HEPT model sampled using approximate MCMC 105

Figure 5.4 Lineage accumulation curves and parameters sampled under the assumption of high (ω = 1.0), intermediate (ω = 0.5), and low (ω = 0.1) extinction heritability 106

Figure 5.5 Posterior predictive simulations under different values of ω..... 107

Figure 5.6 Long-stem phenomenon in simulated trees 108

Figure 5.7 Ratio of the initial waiting time to total tree depth for trees simulated under the ω = 0.1 model..... 109

Figure 6.1 Bayesian consensus phylogram sphenomorphine skinks 127

Figure 6.2 Penalized likelihood chronogram for Australian sphenomorphine skinks indicating taxonomic diversity of each terminal..... 129

Figure 6.3 Stem clade age and extant diversity for Australian skinks 130

Figure 7.1 MCC tree for North American wood-warbler species 149

Figure 7.2 Log-lineage through time plot and distribution of γ for warblers..... 150

Figure 7.3 Maximum likelihood reconstruction of speciation-through-time curve for wood-warblers under overall best-fit model 151

Figure 7.4 Distribution of ΔAIC_{TS} test statistic as tabulated from posterior distribution of wood-warbler phylogenies sampled using MCMC 152

Figure 8.1 Species richness increases with clade age..... 164

Figure 8.2 Correlation between diversity and age for clades simulated under constant rate, relaxed rate, and overdispersed rate models 179

Figure 8.3 Correlation between clade age and species richness under clade volatility simulations 180

Figure 8.4 Distribution of correlations between log(diversity) and age for clades simulated under density-dependent, density-dependent with geographic area, geographic area only, and linear rate change models 181

Figure 8.5 Diversification-through-time curves for the avian tribes dataset..... 182

LIST OF TABLES

Table 1.1 Examples of traits that may cause species selection 11

Table 2.1 Results of fitting four birth-death models to the agamid data 39

Table 2.2 Diversification parameters for the Australian Agamidae under the best-fit model before and after the inferred rate shift 40

Table 4.1 Parameters used for simulating phylogenies undergoing temporal declines in the net diversification rate 78

Table 4.2 Results of fitting constant-rate (birth-death) and variable-rate models to phylogenies of agamid lizards, wood-warblers, and pythons 79

Table 5.1 Run and sampling statistics for approximate MCMC 100

Table 5.2 Gelman-Rubin statistic for parameters across MCMC runs 101

Table 5.3 Diversification parameters for warblers under HEPT model 102

Table 5.3 Relative extinction rates for mean lineage accumulation curves 110

Table 6.1 Analysis of diversification rates in sphenomorphine skinks 131

Table 7.1 Summary of diversification models fitted to the wood-warbler tree..... 153

Table 8.1 Datasets considered in this study and the age-diversity correlation..... 166

Table 8.2 Summary of models fitted to the avian tribes dataset 183

CHAPTER 1

REINVENTING SPECIES SELECTION WITH MOLECULAR PHYLOGENIES

Abstract

Species selection as a potential driver of macroevolutionary trends has been relegated to a largely philosophical position in modern evolutionary biology. Fundamentally, species selection is nothing more than the outcome of heritable differences in speciation and extinction rates among lineages when the causal basis of those rate differences is not determined by genotypic (within-population) fitnesses alone. Here we discuss, in light of species selection theory, the rapidly growing literature on variation in species diversification rates as inferred from molecular phylogenies. We argue that modern studies of diversification rates demonstrate that species selection is an important process influencing both the evolution of biological diversity and distributions of phenotypic traits within higher taxa. Finally, we suggest key questions for future studies.

Introduction

Natural selection is one of the most important ideas in the history of science and philosophy. Charles Darwin, writing in the original edition of the Origin, defined natural selection as "the preservation of favorable variations and the rejection of injurious variation" (1859:81). Although the mechanisms of heredity would remain unknown for decades, Darwin went on to state elsewhere that these 'favorable variations' must be passed from parent to offspring for evolution to occur. A modern definition of selection is "differential survival and reproduction of replicating entities based on heritable trait differences", which – as a statement about process - is fundamentally the same as that proposed by Darwin.

Identifying the levels at which selection acts has been fraught with controversy, and one of the longer-standing controversies in evolutionary biology has been whether selection can occur at the species level (Okasha 2006). To the outsider, the literature on species selection is confusing and rich in philosophical terminology which frequently seems irrelevant to the pertinent evolutionary issues at hand. Here we clarify the concept of species selection and explain why accepting the central premise of species selection is fundamental to understanding large-scale patterns of biological diversity. We believe that a focus on the distinction between species and organismic selection forces us to contrast the determinants of genotypic, within-population fitness with the causes of differential rates of species origination and extinction. We argue that evidence from molecular phylogenies provides a rich demonstration of species selection influencing both patterns of species richness and trait distributions across a range of phylogenetic scales.

What species selection is and is not

Species selection (Stanley 1975) is the outcome of heritable variation in speciation and extinction rates (Jablonski 1986; Grantham 1995). Such variation in diversification thus reflects differences in the emergent fitness of species (Lloyd and Gould 1993; Coyne and Orr 2004), where emergent fitness is simply differential survival and reproduction of species that cannot be extrapolated from analyses of survival and reproduction of lower levels in the hierarchy, such as individuals, social groups, or alleles. Traits under species selection increase the net diversification rate of a lineage (Figure 1.1), where the net diversification rate is the difference between the speciation rate and the extinction rate. A particular trait might cause an increase in the rate of speciation, but there is no way to infer this from consideration of population genetic phenomena alone, because differential survival and reproduction of

individuals within populations has no necessary consequences for species divergence. Even if a trait simultaneously increases fitness at the individual level (by increasing individual survival and reproduction) and the species level (by increasing the speciation rate), the causal mechanisms underlying fitness at these two levels are decoupled.

This is not to argue that organismic selection is unrelated to species selection. Selection at the individual level provides a source of trait variation between species that may result in species selection (Figure 1.1), just as mutation is the origin of variation that results in organismic selection. However, the mechanism by which a trait becomes fixed within a species – whether through selection or drift – is irrelevant to understanding the effects of a trait on diversification. One example of a trait that appears to be favored by both organismic and species selection is floral symmetry. Within populations, plants with more symmetric flowers may experience greater reproductive success than less symmetric individuals, because floral symmetry facilitates pollination efficiency and reduces pollen waste (Armbruster et al. 1994; Kalisz et al. 2006). However, such symmetry incidentally leads to greater pollinator specificity, which in turn facilitates rapid species divergence due to the potential for pollinator isolation among geographically isolated populations; plant clades with symmetric flowers have diversified at greater rates than those with more asymmetric flowers (Sargent 2004). Selection at the individual level presumably led to the evolution of more symmetric flowers within populations, which then served as the raw variation for species selection. The net outcome is that floral symmetry has become more frequent both within populations (organismic selection on pollination efficiency) and within clades (species selection on pollinator specificity).

Some previous authors have argued that species selection can only operate on emergent characters (Vrba 1984), where emergent characters are species traits that

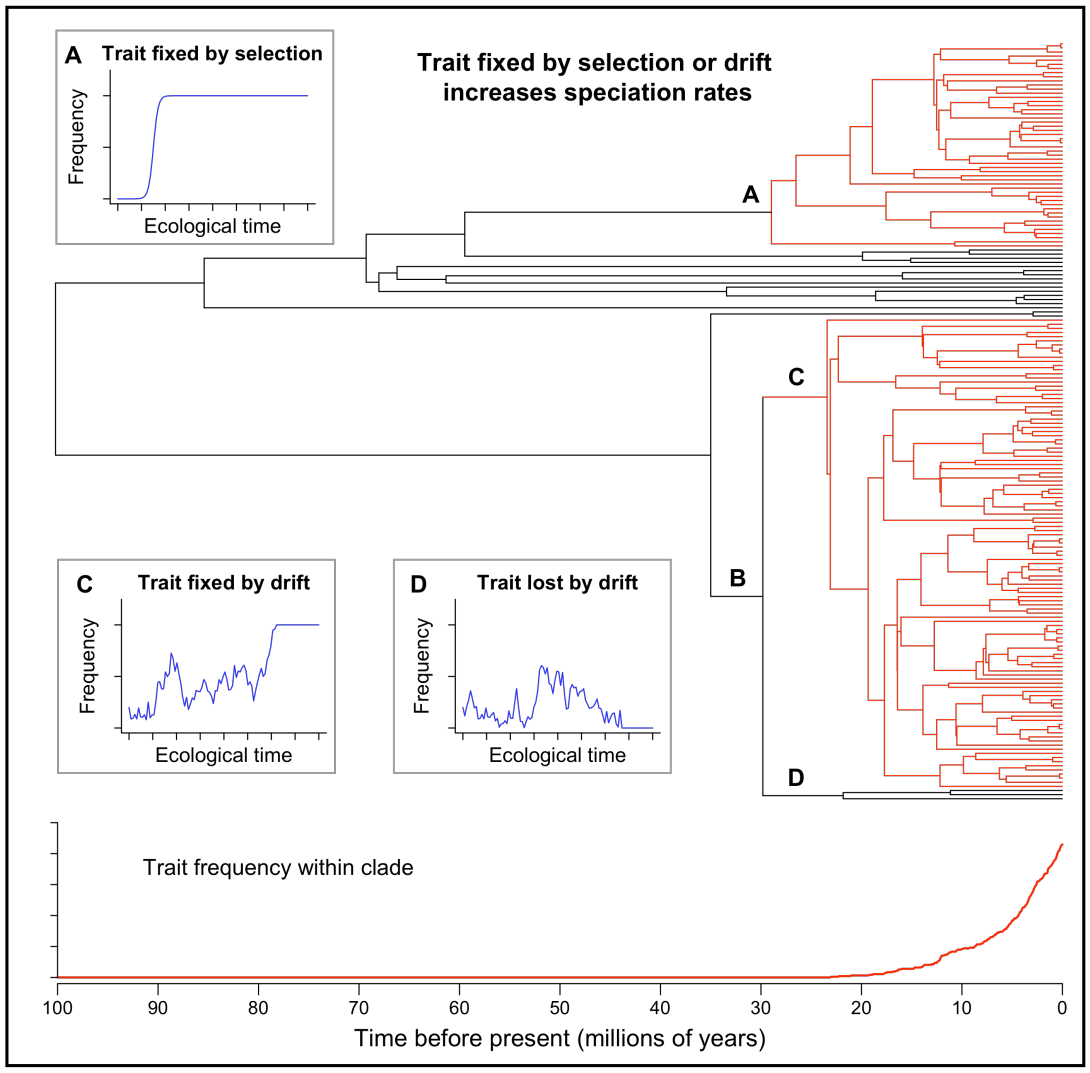

Figure 1. Organismic selection and species selection on a phenotypic trait. A trait causing a five-fold increase in the speciation rate originates via mutation twice in a proliferating clade: in the ancestor of clade A, and again in the ancestor of clade B (lineages with the trait are shown in red). In the lineage leading to clade A, the trait is rapidly fixed by organismic (within-population) selection and its subsequent effect on the speciation rate results in a large increase in diversity of A relative to its sister clade. In the ancestor of clade B, the trait is not favored by within-population selection, but is fixed by drift in one descendant lineage and lost by drift in the other, resulting in a large diversity difference between clades C and D. Timescale of trait fixation by selection or drift is extremely rapid relative to timing of species diversification. Note that the frequency of the trait increases within the clade due to the higher speciation rate of lineages A and C, not through the action of selection within populations.

cannot be reduced to individual traits. Examples of emergent traits include geographic range and genetic population structure, which are not a simple summation of individual trait values (Jablonski 1986, 2008). Aggregate traits, in contrast, can be expressed as the sum of individual trait values. White fur as a trait in polar bears, for example, can be explained as a simple consequence of the fact that all individual polar bears have white fur. While virtually all researchers in the field have agreed that differential speciation and extinction due to emergent traits is species selection (Jablonski 2008), the situation has been more contentious for aggregate characters.

However, it is clear that the effects of both aggregate and emergent traits can influence rates of speciation and extinction in a manner that is entirely decoupled from the effects of those traits at the individual level, as illustrated by the floral symmetry example above (an aggregate trait). There is no absolute or relative within-population fitness value that can – even in principle – be assigned to an individual trait that yields a prediction of that trait's effect on speciation rates (Grantham 1995), and we would argue that this is generally true for extinction rates as well (see below). Fitness values for individual traits predict only changes in trait frequency within a population. Yet, if the trait causes a higher rate of speciation on lineages possessing the trait, it will be present in an increasing percentage of species through time (Figure 1.1). Even if a trait is favored by selection at the levels of both individuals (survival, reproduction) and species (speciation, extinction), our understanding of large-scale evolutionary patterns is contingent on our recognition of distinct levels of selection, because the processes that yield high fitness at the levels of individuals (differential reproduction) and species (differential speciation) with the trait are different.

Some have argued that selective extinction does not necessarily reflect species selection (Grantham 1996; Okasha 2006). It is true that if a species becomes extinct

because of low absolute fitness of all genotypes within the species, then extinction is reducibly specified by population genetic processes. As a simple example, imagine that all individual organisms above a certain threshold body size have low absolute fitness, and that all genotypes within a particular species result in body sizes above the threshold. Individuals thus die because they possess a particular trait value (large body size), and the fact that all individuals die results in the extinction of the species. In this case extinction is arguably due only to organismic selection. However, such a clear case of organismic selection leading to species' extinction is only possible for a very limited set of scenarios: specifically, those where extinction is a direct effect of selection on individual organisms. For such a process to drive a trend, traits under selection at the individual level would have to arise repeatedly, become fixed within populations, and then drive the species to extinction by reducing the absolute fitness of all genotypes to below replacement ($R < 1$).

The evidence that such pure organismal selection directly mediates extinction dynamics is far from convincing. Aggregate traits like body size are usually correlated with emergent traits such as species abundance and geographic range size (McKinney 1997; Purvis et al. 1995), for which selection can only occur at the species level. Moreover, aggregate traits can even be favored at the individual level while at the same time increasing extinction probability (Rankin and Lopez-Sepulcre 2005). For example, organismic selection for resource specialization might incidentally reduce the total population size of a species and thus increase the probability of extinction via chance demographic fluctuations. This would be a clear case of species selection, because of the potential for conflict between individual and species level processes (Okasha 2006).

To summarize, species selection is an important concept because it forces us to recognize that traits can influence patterns of speciation and extinction independent of

their effects on survival and reproduction at the individual level. We cannot even begin to address questions like "why are some groups of organisms so much more diverse than other groups" and "why are some do so many species have sexual reproduction" without accepting that selection can be decoupled at the individual and species level. Rather than debating whether traits are emergent or aggregate, we focus instead on the processes that result in differential diversification.

Molecular phylogenies provide a growing body of evidence for species selection

Although it is not widely acknowledged, questions about variation in diversification rates are fundamentally questions about species selection and the broader role of the process in generating biological diversity. Species selection has in fact been the focus of an ever-expanding research program in comparative biology for nearly two decades, fueled by the availability of both molecular phylogenies and rigorous statistical methods for diversification analysis. At least two patterns should immediately suggest the possibility of species selection. The first is when speciation and extinction rates show heritable differences among lineages (Savolainen et al. 2002). The second involves any repeatable effect of a trait on the rate of diversification of the species possessing the traits (see Figure 2.2). Molecular phylogenies provide a rich source of material with which to assess these patterns. Rates of species diversification can be reconstructed from time-calibrated molecular phylogenies (Mooers & Heard 1997; Rabosky 2006; Ricklefs 2007), and even trees with topological information only can be used to identify shifts in diversification as well as correlations between traits and diversification rates (Mitter et al. 1988).

A substantial literature has developed around methods for testing the relationship between traits and diversification rates (Paradis 2005; Ree 2005; Freckleton et al 2008), but formal tests for the heritability of diversification rates have

been limited (Savolainen et al. 2002; Davies et al. 2004). This might suggest that few tests for heritability of diversification rates have been performed. However, all studies of which we are aware that have documented a relationship between a trait and diversification rate have implicitly demonstrated heritability of rates, whether this relationship was stated or not. Heritability in this context is nothing more than a tendency of progeny species to resemble their 'parents'. Most traits show a tendency to be inherited across speciation events; were this not the case, there would be no need for phylogenetic independent contrasts and other comparative methods. If a trait is associated with increased diversification rates, and the trait is heritable in this sense, diversification rates must themselves be heritable.

Many phylogenetic trees show evidence for dramatic differences in species richness among even closely related clades, a pattern that may be caused by species selection. This is especially true for large phylogenetic trees, which frequently contain a heterogeneous mix of clades with both high and low diversification rates (Magallon and Sanderson 2001; Sims & McConoway 2003; Davies et al. 2004; Hunt et al. 2007). However, geographic area, environmental energy, and other ecological covariates exert a profound effect on patterns of diversification (Davies et al. 2005; Ricklefs 2006; Ricklefs et al. 2007; Rabosky 2009), and these effects are particularly apparent at larger phylogenetic scales. Here, a clade may appear to have a high (heritable) net rate of diversification because of geological or climatic phenomena specific to the location of the clade (Cracraft 1982). Such biogeographic processes might cause increased rates of diversification (Moore and Donoghue 2007) or impose ecological limits on clade growth (Ricklefs 2007; Rabosky 2009).

Yet these extrinsic influences on diversification are frequently entangled with lineage-specific factors that may promote speciation. This is one explanation for the observation that diversification of different groups which inhabit the same geographic

and ecological theatre frequently leads to radically different evolutionary outcomes. For example, Hawaii's honeycreeper finches have undergone a spectacular evolutionary radiation, while the remaining four passerine bird lineages in the islands have not (Lovette et al. 2002). These finches did not arrive earlier on the islands than other groups, and their high rate of diversification has been attributed to evolutionary lability of feeding morphology (Lovette et al. 2002). Similar arguments have been made to explain the radiation of Darwin's finches and the failure of other clades on the Galapagos to radiate (Grant and Grant 2007). In Australia, a single clade of arid-adapted scincid lizards has diversified into nearly 200 species and is far more diverse than all other Australian skink clades (Rabosky et al. 2007), even though several clades that did not radiate occupy similar ecological settings. This suggests that intrinsic traits, perhaps associated with thermal physiology, have been a key factor in the explosive diversification of the group. Cichlid fishes have undergone dramatic evolution radiations in each of the major Rift Valley lakes, but comparatively little diversification has occurred within the other major teleost fish clades in the lakes (Greenwood 1984). Clearly, we have much to learn about the intrinsic factors that promote diversification in these and other groups.

Studies of species selection are most informative when the causal basis of differential diversification can be identified (Figure 1.2). A variety of traits are known to result in differential diversification (Coyne & Orr 2004; Jablonski 2008) (Table 1.1). Among the first examples of traits to show putative correlations with diversification rates were those associated with sexual selection (Barraclough et al. 1995). Under this model, species that are prone to rapid divergence in traits that influence mating success will, all things being equal, be more likely to undergo speciation than those showing low potential for divergence in such traits. This is likely

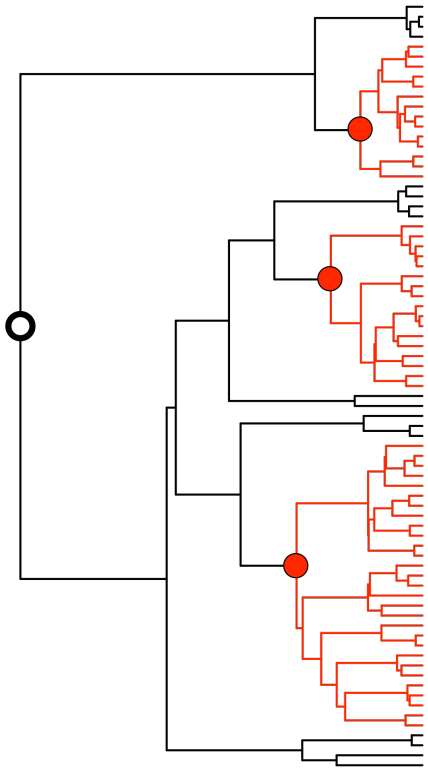

Figure 2. Sister clade contrasts reveal effects of a trait on diversification. Lineages acquiring a derived state of a trait (red circle) from an ancestral state (open black circle) undergo a threefold increase in the speciation rate. As a result, they are consistently more diverse than their sister clades. "Sister clade contrasts" are widely used to test for the effects of such traits on diversification when multiple pairs of sister clades can be identified that differ with respect to a trait of interest. Sister clades are by definition the same age, and any differences in diversity are due to chance or to real differences in diversification rates. Under the null hypothesis that a trait has no effect on diversification, we expect clades without the trait to be, on average, as diverse as clades with the trait. These data can be analyzed with a simple binomial test. In this example, we have three pairs of sister clades, and the clade with the trait is more diverse in each case than its sister taxon. The probability that there is no effect of the trait on diversification is $p = 0.125$. In this hypothetical case, each instance of 'red' trait evolution results in a diversity increase, but our small sample size gives prohibitively low power (a minimum of five clades would be needed to see a significant effect at $p < 0.05$).

Table 1.1. Examples of traits that may cause species selection

Trait	*Possible causes of differential diversification*	*References*
Intensity of sexual selection	Geographically isolated populations quickly diverge in sexually selected characters and reciprocally fail to recognize members of the other population as potential mates following secondary contact.	Barraclough et al. 1995; Coyne & Orr 2004; Price 2008
Population density	Decreased risk of extinction due to stochastic fluctuation in population size	Kruger 2008
Annual dispersal	New populations (incipient species) founded at greater rates with increased dispersal	Phillimore et al. 2006
Limb complexity	Evolution of limb complexity may facilitate divergence into new ecological niches; may permit evolution of more complex sexual display and copulation behavior, leading to increased sexual selection (see above).	Adamowicz et al. 2008

due to the fact that traits under sexual selection serve an incidental role in species recognition: geographically isolated populations that diverge in sexually selected traits may no longer recognize each other as potential mates following secondary contact (Price 2008). Many studies have reported positive effects of such traits on diversification rates, in groups that include agamid lizards (Stuart-Fox et al. 2003), birds (Seddon et al. 2008; Kruger 2008), flies (Katzourakis 2006), and fishes (Mank 2007). The evidence from birds is mixed, with other studies reporting no relationship between sexual selection and diversification (Morrow et al. 2003; Phillimore et al. 2006). However, Price argues that the lack of effect may be attributable to the fact that sexual selection is so widespread within birds (Price 2008), suggesting that other factors might be more important influences on species richness.

In addition to floral symmetry, a number of other traits associated with pollinator specificity have been shown to correlate with diversification rates. For example, lineages with nectar spurs have higher diversification rates than unspurred lineages, across multiple clades of plants (Hodges 1997; Kay et al. 2006). Nectar spurs restrict the pool of possible pollinator species and are hypothesized to facilitate more rapid evolution of reproductive isolation among geographically isolated populations. However, at the individual level, nectar spurs appear to evolve because of selection for improved pollen transfer (Whittall & Hodges 2007); their effects on speciation rates are entirely incidental with respect to their consequences for within-population fitness. Other floral characters that should increase pollinator specificity appear to show similar effects, including biotic (animal) pollination (Dodd et al. 1999; Kay et al. 2006).

We do not intend to be exhaustive, but numerous other traits show positive (or negative) effects on diversification rates, ranging from behavioral flexibility in birds (Nicolakakis et al. 2003) to the evolution of a climbing habit in plants (Gianoli 2003).

Even the neutral rate of molecular evolution may be associated with elevated diversification (Barraclough & Savolainen 2001; Webster et al. 2003).

Challenges in the study of species selection

The central challenges faced by these and related studies are fourfold. First, almost all 'traits' are in fact highly correlated with a suite of traits that confound simple interpretations of patterns. This is amply demonstrated by analyses of traits associated with extinction risk in extant species (Purvis et al. 2005), which are so inseparably bound with other traits that it is difficult to assign causation. Second, because numerous traits can influence diversification rates, traits may oppose one another at a given level of the selection hierarchy. The complexity of these interactions may explain, in part, why traits show contrasting effects on diversification among even closely related groups of higher taxa (Vamosi and Vamosi 2004; Isaac et al. 2005).

A third concern is strictly methodological: with some simple methods of analysis, it can be difficult to distinguish between the repeated effects of a trait on diversification and asymmetric rates of character change (Maddison 2006). In particular, simple sister clade contrasts (Mitter et al. 1988) can lead one to infer that diversification rates differ with respect to a particular trait, when the pattern is solely attributable to the characters themselves evolving at different rates. If one character state evolves frequently, but reversals to the previous state occur rarely, then simple contrasts will reveal an apparent association between that character state and the diversification rate. This may be less problematic for newer methods that use all of the information in phylogenetic trees (Paradis 2005, 2008; Maddison et al. 2007; Freckleton et al. 2008), including branch length data, but researchers need to bear this potential pitfall in mind when interpreting diversification analyses. Finally, that we

can ascribe a given pattern of diversification to a particular trait tells us nothing about the actual mechanisms by which the trait influences speciation and/or extinction dynamics. We argue in favor of a species selectionist viewpoint precisely because we feel that this level of analysis focuses attention on identifying these mechanisms.

While we may have strong reasons to believe that the above examples represent species selection, we still do not know in most cases whether the effects are driven by variation in speciation or extinction rates. Disentangling these processes is an important step towards understanding causation. Because speciation and extinction leave different signatures in phylogenies of extant taxa only (Nee et al. 1994; Rabosky & Lovette 2008), it may be possible to determine whether increased diversification rates are mediated by increased speciation, decreased extinction, or both. The binary state speciation-extinction model (BiSSE) of Maddison et al. (2007) has considerable promise in this regard and enables simultaneous estimates of speciation, extinction, and transition rates for a binary character on a molecular phylogenetic tree.

Also promising is the use of theory to derive a priori predictions for the effects of traits on individual and species fitness. This approach has been used to explore the evolution of specialist versus generalist pollination (Sargent & Otto 2006) as well as the evolution of sexually dimorphic floral displays (Vamosi & Otto 2002). Experimental studies of the evolution of reproductive isolation may also help refine interpretations of higher-level patterns. For example, Arnqvist et al. (2000) reported increased diversification rates in clades of insects where females mate with multiple males relative to those where females mate only once, suggesting that postmating sexual selection might have caused differential diversification. But several recent experimental studies on Drosophila found no evidence that experimental manipulation of postmating sexual selection led to more rapid evolution of reproductive isolation. This suggests that postmating sexual selection influences extinction but not speciation

rates, or alternatively that other, correlated traits have caused differential diversification in insect clades with contrasting mating systems.

Prospects and conclusions

Whether species selection can and does occur is, in our minds, an issue settled. However, many issues remain, in addition to those outlined above. We consider the following to be of paramount importance. First, how much of the variance in species richness among clades can be explained by species selection relative to extrinsic influences on diversification? These explanations are not mutually exclusive and there is clear evidence for interactions between intrinsic and extrinsic factors (Lovette et al. 2002; Rabosky et al. 2007; Jablonski 2008). The answer to this question is almost certainly complex, but at present, our understanding of these factors and their interactions is deeply unsatisfying. Second, could species selection drive large-scale trends in organismal design? It is possible that some of the most striking trends in evolution may be attributable in part to species selection, such as the evolution of phenotypic complexity (Adamowicz et al. 2008) and even the capacity of species to evolve (Pigliucci 2008; Gerhart & Kirschner 1998). Finally, how are human activities interacting with species traits to influence extinction rates? Species traits with high heritability currently explain much of the variation in extinction risk among threatened species (Purvis et al. 2005; Vamosi & Vamosi 2008), leading Jablonski to suggest that we may be unwittingly be conducting a massive experiment in species selection (Jablonski 2008). An explicit recognition of selection at the species level is central to our ability to answer these and other questions, because it focuses our attention on the actual mechanisms by which traits influence the dynamics of speciation and extinction.

CHAPTER 2

LIKELIHOOD METHODS FOR DETECTING TEMPORAL SHIFTS IN DIVERSIFICATION RATES

Abstract

Maximum likelihood is a potentially powerful approach for investigating the tempo of diversification using molecular phylogenetic data. Likelihood methods distinguish between rate-constant and rate-variable models of diversification by fitting birth-death models to phylogenetic data. Because model selection in this context is a test of the hypothesis that diversification rates have changed over time, strategies for selecting best-fit models must minimize Type I error rates while retaining power to detect rate variation when it is present. Here I examine model selection, parameter estimation, and power to reject the null hypothesis using likelihood models based on the birth-death process. The Akaike Information Criterion (AIC) has often been used to select amongst diversification models; however, I find that selecting models based on the lowest AIC score leads to a dramatic inflation of the Type I error rate. When appropriately corrected to reduce Type I error rates, the birth-death likelihood approach performs as well or better than the widely used gamma statistic. Analysis of datasets simulated under a range of rate-variable diversification scenarios indicates that the method has much greater power to detect variation in diversification rates when extinction is present. Furthermore, the birth-death likelihood method appears to be the only approach available that can distinguish between a temporal increase in diversification rates and a rate-constant model with non-zero extinction. I illustrate use of the method by analyzing a published phylogeny for Australian agamid lizards.

Introduction

There is currently great interest in understanding variation in speciation and extinction rates both over time (e.g., Hey 1992; Sanderson and Donoghue 1996; Barraclough and Nee 2001; Zink et al. 2004; Ruber and Zardoya 2005) and among lineages (e.g., Slowinksi and Guyer 1989; Mooers and Heard 1997; Coyne and Orr 2004; Ree 2005). The proliferation of molecular phylogenetic data has provided researchers in this area with a wealth of material for comparative analyses of diversification. Many intriguing and unanswered questions in evolutionary biology concern temporal variation in diversification rates. For example, how have climatic shifts influenced the tempo of diversification (Zink and Slowinski 1995; Kadereit et al. 2004; Weir and Schluter 2004)? How do the dynamics of speciation and extinction differ during evolutionary radiations on islands and in continental systems (Losos and Schluter 2000; Harmon et al. 2003)? Is early, rapid diversification a general feature of adaptive radiation (Lovette and Bermingham 1999; Schluter 2000)?

To address such questions, researchers utilize an increasingly sophisticated statistical toolkit. The framework for analyzing temporal variation in diversification rates typically entails comparing observed patterns of diversification to those generated by an appropriate null model where rates have been constant over time. The most widely used null model for diversification rate analyses is the birth-death model (Kendall 1948), where clades grow under a constant per-capita speciation and extinction rate. A popular approach is to test observed lineage diversity through time against rates of lineage accumulation under a rate-constant model (Nee et al. 1992; Wollenberg et al. 1997). Both parametric and non-parametric statistics have been developed to test the distribution of internal nodes in a molecular phylogeny for departures from the pure birth variant of the birth-death process (Zink and Slowinski 1995; Paradis 1998b; Pybus and Harvey 2000).

Likelihood methods based on the birth-death model are a potentially powerful approach for reconstructing the history of diversification from phylogenetic data. In contrast to other methods, such model-based approaches can simultaneously assess rate variation over time and provide estimates of relevant diversification parameters (Sanderson 1994; Barraclough and Vogler 2002). Likelihood is increasingly used to reconstruct shifts in diversification rates, either through explicit use of the birth-death model (Turgeon et al. 2005) or through a related approach involving survival analysis (Paradis 1997; Emerson et al. 2000; Pitra et al. 2004; Paradis 2005).

While the array of statistical methods for diversification rate analyses continues to grow, there have been few critical assessments of the performance of likelihood and other methods against datasets simulated under both rate-variable and rate-constant diversification scenarios. There are several important issues that should be considered by researchers wishing to make use of these and other methods.

Although model selection is the subject of a substantial and growing literature in molecular phylogenetics (Zhang 1999; Posada and Crandall 2001; Pol 2004; Posada and Buckley 2004), little attention has been given to methods for selecting amongst rate-constant and rate-variable models of diversification. Because researchers typically use model selection in this context as a test of the hypothesis that diversification rates have varied over time, it is imperative that model selection be based on non-arbitrary criteria that minimize the probability of erroneously rejecting a true rate-constant model of diversification in favor of a model where rates have varied over time (Type I error rate).

A second issue concerns the relative power of diversification rate test statistics. In particular, we might ask whether it is ever possible to detect a temporal increase in diversification rates using phylogenies of extant taxa only, because both increased diversification rates and constant, non-zero extinction result in an apparent excess of

recently diverged lineages (Nee et al. 1994a; Kubo and Iwasa 1995). Some have questioned whether any method failing to explicitly consider the confounding effects of extinction could detect an increase in diversification rates (Nee 2001).

In this paper, I address model selection and Type I error rates for likelihood-based analyses of diversification. I evaluate the power of the birth-death model to recover the "true" model of evolution when rates are known to have varied over time, and I compare the birth-death approach to the γ-statistic (Pybus and Harvey 2000), a popular parametric alternative. Because we know little about bias and variances of parameters estimated under the birth-death model when rates are known to have varied over time, I examine the relative error of parameters inferred from phylogenies simulated under several models of diversification. Finally, I illustrate use of the likelihood approach by analyzing the tempo of diversification in a radiation of Australian lizards.

Methods

Models for Likelihood-Based Analyses of Diversification

Here I consider the generalized birth-death process as a framework for the analysis of diversification rates. This is a simple null model for the growth of a phylogenetic tree over time: existing lineages give birth to new lineages at a per-lineage rate λ and go extinct at a per-lineage rate μ. The Yule process or pure-birth model (Yule 1924) is a special case of the birth-death process with $\mu = 0$, where the number of lineages can only increase over time. In contrast, when $\mu > 0$, the number of lineages can decrease over time; a clade diversifying under this model can become extinct even when the speciation rate exceeds the extinction rate.

Kendall (1948) formalized the probability that a stochastic process beginning with n0 lineages will have nt progeny after some time t under any constant and time-

dependent birth and death rates. Nee et al. (1994b) extended the Kendall model to the case of reconstructed phylogenies; in this case, birth events (speciation) can only be observed if each descendent lineage leaves at least one surviving progeny in the present (Figure 2.1).

The probability of each waiting time between successive speciation events is equal to the probability that each of the n lineages at the start of the interval has exactly one progeny (itself) after some time t, conditioned on the probability that none of the n lineages go extinct (Nee et al. 1994b; eqn. 17):

(2.1) Prob (t | n, λ, μ, T, tn) =

$$n(\lambda-\mu)\exp(-(\lambda-\mu)nt)\frac{\left(1-\frac{\mu}{\lambda}\exp\left[-(\lambda-\mu)(T-t_n-t)\right]\right)^{n-1}}{\left(1-\frac{\mu}{\lambda}\exp\left[-(\lambda-\mu)(T-t_n)\right]\right)^{n}}$$

where T is the time from the root node to the present, tn is the birth time of the nth lineage, and t corresponds to the waiting time until the birth of another lineage that can be observed in a reconstructed phylogeny.

From (2), we multiply the transition probabilities together to obtain the likelihood that a particular birth-death process has produced the branching times observed in the reconstructed phylogeny (Nee et al. 1994b). Reparameterizing the model such that r = λ – μ (the net diversification rate), a = μ / λ (the extinction fraction), and letting x be a vector of observed branching times (Figure 2.1), we have

(2.2) $L(\mathbf{x} \mid a, r) =$

$$\prod_{n=2}^{N-1} n\, r \exp\left(-n\, r\left(x_n - x_{n+1}\right)\right) \frac{\left(1 - a \exp\left[-r\left(x_{n+1}\right)\right]\right)^{n-1}}{\left(1 - a \exp\left[-r\left(x_n\right)\right]\right)^{n}}$$

Here I consider the special case where a clade diversifies under parameters r1 and a1, until some point in time ts, where rates shift to r2 and a2. Due to the computational time required to optimize parameters over all possible values of ts, I consider only observed branching times as possible shift points. This is a conservative approach, because consideration of shift points other than the observed branching times can only increase the likelihood of the rate-variable model relative to the rate-constant model. This approach has been implemented in several previous studies (Barraclough and Vogler 2002; Turgeon et al. 2005). Throughout the text, I refer to this method as BDL.

An alternative likelihood-based approach to testing for temporal shifts in diversification rate is survival analysis (Paradis 1997, 1998). Here we think of the time-axis of the phylogeny in reverse: a speciation event in the reconstructed phylogeny becomes a failure event in survival analysis. Each lineage has a probability of h(t) of failure or death; in the case of a reconstructed phylogeny with exponentially distributed branching times, h(t) is equal to the net diversification rate (Cox and Oakes 1984; Paradis 1997). Survival analysis has been used to test for temporal increases in net diversification rates (Near et al. 2003); however, because survival models do not include an extinction term, it is unclear whether they can separate this phenomenon from constant background extinction rates.

I used the birth-death model to analyze phylogenies simulated under a set of rate-variable and rate-constant models of speciation and extinction. Likelihoods of simulated phylogenies were computed under two rate-constant and two rate-variable

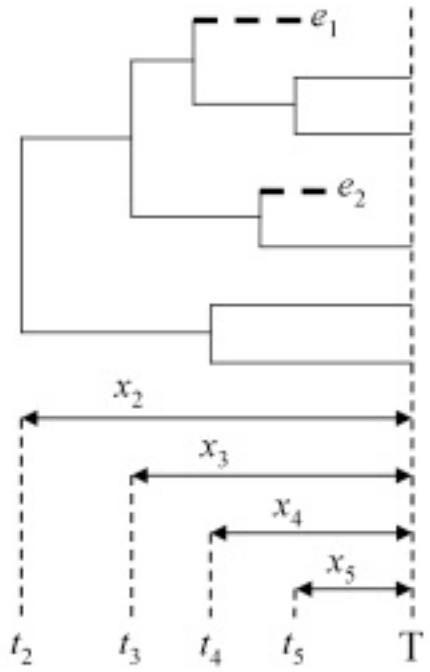

Figure 2.1. A reconstructed phylogenetic tree with five species to illustrate parameters discussed in text. T represents the total time elapsed from root node to the present, t_i is the birth time of the i^{th} lineage, and x_i is the branching time of lineage i ($x_i = T - t_i$). Lineages e_1 and e_2 went extinct at some time before the present, and we are thus unable to see the corresponding speciation events in a reconstructed phylogeny.

models: 1) the pure birth model (one parameter, λ, with μ set to zero); 2) a constant-rate birth-death model (two parameters, $\lambda > 0$; $\mu \geq 0$); 3) a pure birth rate-variable model where speciation rate $\lambda 1$ shifts to rate $\lambda 2$ at some time ts (three parameters: $\lambda 1$, $\lambda 2 > 0$; $0 < ts < T$); and 4) a rate-variable model with two speciation rates and two extinction rates, but constrained such that the extinction fraction μ / λ remains constant (four parameters: $\lambda 1, \lambda 2 > 0$; $\mu 1, \mu 2 \geq 0$; $0 < ts < T$; but $\mu 1 / \lambda 1 = \mu 2 / \lambda 2$). The distribution of branching times in a reconstructed phylogeny is a function of the net diversification rate, $\lambda - \mu$, and the extinction fraction, μ / λ, and throughout the text I denote these parameters by $r = \lambda - \mu$ and $a = \mu / \lambda$. I also computed likelihoods of rate-constant phylogenies using survival models to assess Type I error rates under different values of a.

For each simulated phylogeny, parameters were optimized for each birth-death model under consideration. I then calculated the likelihood of the data given the model for all models using these optimized parameters. For survival analysis, the likelihood of the simulated data was calculated under two models: a rate-constant model, and a rate-variable model where the diversification rate is modeled as a function of time, $r(t) = \alpha\beta(\alpha t)\ \beta - 1$. In the latter case, the distribution of failure times follows a Weibull distribution, where the diversification rate decreases over time if $\beta > 1$, increases if $\beta < 1$, and is constant if $\beta = 1$.

Model Selection

In the likelihood methods discussed here, model selection is equivalent to hypothesis testing: we are asking, at least initially, whether we can reject a rate-constant model in favor of a model where diversification rates have varied over time. The Akaike Information Criterion (Akaike 1973), or AIC, has been widely used to select amongst different models of diversification (e.g., Paradis 1997; Emerson et al.

2000; Barraclough and Vogler 2002; Pitra et al. 2004). Typically, AICs are calculated for a set of models, and the model with the lowest AIC is taken to be the model that best approximates the data. The AIC is a function of both the log likelihood and the number of free parameters in a given model:

(2.3) $AIC = -2 \log L + 2p,$

where p is the number of parameters that are estimated from the data. Thus, more parameters typically improve the fit of the model to the data, but also increase the penalty term (2p).

Some argue that it is inappropriate to use the AIC or any other information-theoretic measure for hypothesis testing, in part because it is believed that the traditional hypothesis testing framework is uninformative and based on arbitrary rejection criteria (Burnham and Anderson 2002). A philosophical discussion of this matter is beyond the scope of the present paper, but several considerations justify the use of the AIC for the present purpose: (1) virtually all research in this area has employed the AIC as a test statistic to distinguish between rate-constant and rate-variable models of diversification; and (2) as will be shown, the AIC performs well in practice and lends itself easily to Monte Carlo methods used to infer the distributions of other parametric and non-parametric test-statistics.

Is it sufficient to select the model with the lowest AIC, regardless of the difference in AIC scores between the best and second-best models (Paradis 1997)? This issue is essentially one of confidence: if the best rate-variable model represents an improvement in fit of x AIC units over the best rate-constant model, how confident can we be that the rate-variable model is the better approximation of our data?

To address this issue, I simulated 1000 phylogenies each of N = 50 taxa under ten rate-constant diversification scenarios. The net diversification rate r was identical for all ten scenarios, but the extinction fraction a was varied from a = 0 to a = 0.9. The likelihood of each simulated phylogeny was calculated under the four variants of the birth-death model and two survival models described previously. AIC scores were computed for each model to address the following questions: (1) Is the birth-death likelihood approach susceptible to high Type I error rates? (2) Can objective criteria based on the AIC be established to select amongst rate-variable and rate-constant models of diversification such that the Type I error rate is minimized?

For each simulated phylogeny, the difference in AIC score between the best rate-constant and rate-variable models was calculated as

(2.4) $\Delta AICRC = AICRC - AICRV$

where AICRV is the lowest AIC score among rate-variable models (hereafter referred to as the candidate set) and AICRC is the lowest AIC score among the two rate-constant models. $\Delta AICRC$ is positive if a rate-variable model best fits the data and negative if a rate-constant model is the better fit. Use of $\Delta AICRC$ as a test-statistic permits us to identify a difference in AIC scores between rate-constant and rate-variable models such that $\alpha \pounds 0.05$. The $\Delta AICRC$ giving a Type I error rate of $\alpha = 0.05$ corresponds to the 95^{th} percentile of the distribution of $\Delta AICRC$ scores tabulated from phylogenies simulated under the null hypothesis of rate-constancy.

One might predict that the probability of Type I error will increase for larger phylogenies, because of the greater number of likelihood estimates (under each rate-variable model, we have a likelihood of a rate shift for each branching time). If this is the case, AIC criteria for rejecting the null hypothesis of rate-constancy will require

adjustment based on the number of taxa in our tree. I simulated 1000 phylogenies of N=15, N = 30, N = 60, and N = 100 taxa under the pure birth model to determine whether Type I error rates show dependency on sample size.

It is also possible that the Type I error rate will increase as the number of rate-variable models under consideration increases. For each set of simulated phylogenies, I computed the distribution of ΔAICRC under three different candidate sets of rate-variable models. In addition to the two rate-variable models described previously, I added a five parameter variant of the pure birth model with three speciation rates and two shift points, all of which were optimized for each simulated phylogeny. This model may be useful in many situations of interest to biologists; for example, rapid diversification early in the history of a clade could limit our power to detect recent rate shifts that may have occurred (for example) during the Pleistocene. By decoupling these processes, we may better be able to approximate the true tempo of diversification. The candidate sets of rate-variable models for this analysis consisted of (1) the three parameter model only; (2) three and four parameter models; and (3) three, four and five parameter models.

Finally, I used survival analysis to assess the distribution of ΔAICRC as a function of the extinction fraction a. If there is a positive relationship between ΔAICRC and a, this method cannot be used to distinguish temporal increases in diversification from a rate-constant model with $a > 0$.

Power to Detect Temporal Variation in Diversification Rates

A good statistical framework for diversification rates analysis cannot simply minimize Type I error: it must have sufficient power to detect temporal variation in diversification rates when it is present. Of primary concern is whether we can ever detect a temporal increase in the net diversification rate using phylogenies of extant

taxa only. Constant, non-zero extinction rates produce an apparent excess of recently diverged lineages relative to the pure-birth model (Nee et al. 1994a; Kubo and Iwasa 1996), possibly leading to the erroneous conclusion that diversification rates have increased over time (Nee 2001). The most widely used diversification rates test-statistic, γ, can only be used to detect temporal decreases in diversification (Pybus and Harvey 2000). This fact is not widely appreciated, and positive γ values have been interpreted to support temporal increases in diversification rates (e.g., Linder et al. 2003; Turgeon et al. 2005).

I assessed the power of the birth-death likelihood approach by simulating 500 phylogenies of N = 30, N = 60, and N = 100 taxa under each of eight models of diversification, yielding a total of 24 sets of 500 phylogenies. For each N, phylogenies were simulated with a constant extinction fraction a = 0 or a = 0.5 under the following scenarios: 1) five-fold decrease in the net diversification rate r; 2) two-fold decrease in r; 3) two-fold increase in r; and 4) five-fold increase in r. Clades grew under rate r1 from an initial size of two lineages until the birth of the N/2 lineage. Rates then shifted to a second diversification rate r2, and the clade was permitted to grow until the Nth lineage was born. Thus, a phylogeny of 60 taxa would grow under rate r1 until the birth of the 30^{th} lineage, after which the clade continued to grow with a new rate r2 until the clade reached a final size of N = 60. Extinction was maintained at a constant level of a = 0 or a = 0.5 throughout the duration of the simulation.

The power of the likelihood method to recover the "true" rate-variable model of evolution was simply the percentage of trees for which the null hypothesis of rate-constancy was rejected. The rate constant model was rejected only if ΔAICRC was less than the critical value of the simulated null distribution. The likelihood of each phylogeny was assessed under two rate-constant and two rate-variable models of diversification.

I contrasted the power of the likelihood approach to that of the γ-statistic, a test of the distribution of internal nodes in a phylogeny. Because the γ-statistic tests for departures from the pure birth (a = 0) model of diversification, it cannot distinguish between an increase in speciation and a constant rate model with a > 0 (Pybus and Harvey 2000); both scenarios result in an apparent excess of speciation events near the present relative to the pure birth model. The γ-statistic follows a standard normal distribution under the pure birth process, and we reject the null hypothesis of rate-constancy in favor of a temporal decrease in diversification when $\gamma < -1.645$.

I further explored the power of these methods to detect temporal decreases in diversification rates when rates have shifted earlier or later than the N/2 speciation event. I simulated a two-rate branching process where the net diversification rate decreased 3.5-fold with the birth of the kth taxon (k = 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, or 55) under background extinction rates of a = 0 and a = 0.5. I generated 500 phylogenies of N = 60 taxa for each combination of a and k and applied both BDL and the γ-statistic to each set of simulated phylogenies.

In the above rate-variable scenarios, diversification rates shift instantaneously from rate r1 to r2. Because the likelihood models implemented here are designed precisely to detect this type of rate shift, it is possible that the γ-statistic would perform better than likelihood methods under models where diversification rates have changed gradually over time. To address this possibility, I considered a model of density-dependent cladogenesis, where diversification rates are inversely related to the number of lineages surviving at any point in time. I simulated two sets of 500 phylogenies (N = 60 taxa; a = 0 or a = 0.5) under a density-dependent model with $r(n) = r0\ n^{-x}$, with x = 0.45. Likelihoods were computed under each model as described above.

Parameter Estimation

A potential advantage of the likelihood approach is that it provides estimates of the timing and magnitude of rate shifts. It is not possible to estimate extinction rates from molecular phylogenies with any reasonable degree of confidence, but estimates of the net diversification rate fare much better (Nee et al. 1994a; Kubo and Iwasa 1995; Paradis 2004). To date, no studies have assessed the performance of the BDL method in parameter estimation when rates are known to have varied over time. For each set of N = 30, N = 60, and N = 100 phylogenies simulated under four different rate-variable models, I asked whether likelihood methods could reasonably approximate the net diversification rates before and after the rate shift, r1 and r2, and the timing of the rate shift, ts.

Because phylogenies were simulated under a specified model of diversification, I calculated the relative error in the estimates of r1 and r2 as

(2.5) $$\frac{\hat{r}_i - r_i}{r_i}$$

where is estimated from the data and ri is the true value. Thus, positive values indicate overestimates of the net diversification rate, and negative values indicate underestimates. I expressed the error in ts as a percentage of the total age of each simulated phylogeny,

(2.6) $$\frac{\hat{t}_s - t_s}{T}$$

where ts(hat) is the inferred time of the rate shift in time units from the start of the simulation (t = 0), ts is the true shift time, and T is age of the simulated phylogeny.

Thus, a positive value indicates that the inferred rate shift is later than the true shift point.

Example: Australian Agamid Radiation

To illustrate how these methods might be used, I analyzed a recent molecular phylogeny for Australian lizards in the family Agamidae (Harmon et al. 2003). The agamids ("dragon lizards") constitute a substantial component of squamate reptile diversity in Australia and show considerable diversity in morphology and ecology (Pianka 1986; Melville et al. 2001). The analysis of Harmon et al. (2003) was based on 69 extant taxa, including at least 93% of the species known from Australia.

The agamid tree was constructed by maximum likelihood from 1800 bp of mitochondrial DNA using the GTR + I + Γ model of sequence evolution (Harmon et al. 2003), and the tree was made ultrametric using non-parametric rate smoothing (NPRS; Sanderson 1997). I evaluated the tempo of diversification in the Agamidae with BDL, fitting two rate-constant and two rate-variable models of speciation and extinction using maximum likelihood methods described in this paper. To infer diversification parameters, an age of 30 million years was assigned to the basal divergence between Australian and southeast Asian agamids (Hugall and Lee 2004).

Results

Model Selection

When data are simulated under a model where rates do not vary over time, selecting the model with the lowest AIC score leads to Type I error rates exceeding 28% in all cases (Figure 2.2A); for some values of a, the error rate exceeded 40%. For phylogenies simulated under rate-constant models, the 95th percentile of the distribution of ΔAICRC corresponds to $\alpha = 0.05$: by definition, 5% of the differences

in AIC scores between the best rate-constant and rate-variable models exceed this value. The 95th percentile values are approximately constant or decreasing slightly from a = 0 to a = 0.9 (Figure 2.2B).

There is a high variance in ΔAICRC values, accounting for the ragged distribution of error rates and α0.05 rejection regions. Because error rates do not increase with extinction levels, we can reject the rate-constant model if ΔAICRC for a test phylogeny is greater than the critical value determined by simulating rate-constant phylogenies under the pure-birth process. The lack of a positive relationship between a and ΔAICRC suggests that the birth-death likelihood approach may be capable of detecting a temporal increase in the net diversification rate.

The ΔAICRC required to maintain α £ 0.05 increases with sample size (Figure 2.2C): as the number of taxa increases, we require greater differences in likelihood scores between the best rate-constant and rate-variable models to reject the null hypothesis. When N is small, there is a high variance in the 95th percentile values of the ΔAICRC distribution, but this decreases for larger phylogenies. For phylogenies of N = 60 and N = 100 taxa, Type I error rates exceeded 50% if the model with the lowest AIC is selected as that which best approximates the data. We have little confidence that a rate-variable model best approximates the data when the AIC difference between rate-constant and rate-variable models is less than 3. Only when ΔAICRC approaches 4 for small (N = 30) phylogenies and 5.5 for large (N = 100) phylogenies can the rate-constant model be rejected with confidence.

There is a positive relationship between ΔAICRC required to maintain α £ 0.05 and the number of rate-variable models under consideration (Figure 2.2D). As more complex models are added to the candidate set, Type I error rates increase unless ΔAICRC rejection criteria are adjusted accordingly. This is particularly apparent with

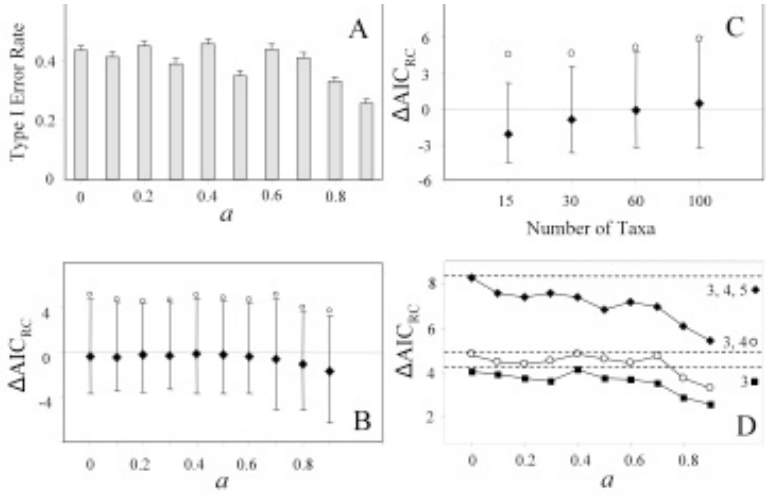

Figure 2.2. Type I error rates and ΔAIC_{RC} rejection criteria for rate-constant phylogenies simulated under different values of the extinction fraction *a* and analyzed with a set of rate-constant and rate-variable birth-death models. (A) Type I error rate as a function of *a* if the model with the lowest AIC score is selected as that which best approximates the data. Type I error rates range between 28-45%, but do not increase substantially with *a*. (B) Differences in AIC scores between the best rate-constant and rate-variable models as a function of *a*. Shown are median ΔAIC_{RC} scores (black diamonds), 0 - 95th percentile range of ΔAIC_{RC} (error bars), and bootstrap standard errors of the 95th percentile ΔAIC_{RC} value (open circles). The 95th percentile of the ΔAIC_{RC} distribution (upper-bound error bar) equals the AIC difference between rate-constant and rate-variable models corresponding to $\alpha = 0.05$ and does not increase with *a*. Negative values of ΔAIC_{RC} are obtained when AIC scores for the best rate-variable model exceed those of the best rate-constant model. (C) Distribution of ΔAIC_{RC} as a function of the number of taxa in simulated phylogenies: as the number of taxa increases, a greater difference in AIC scores between the best rate-constant and rate-variable models is required to maintain $\alpha = 0.05$. (D) 95th percentile of the distribution of ΔAIC_{RC} when data are analyzed with multiple rate-variable models. Dashed lines indicate ΔAIC_{RC} rejection criteria required to maintain $\alpha = 0.05$ when analysis is based on different sets of 3, 4, and 5 parameter rate-variable models (model set indicated adjacent to each dashed line). As more parameters are added to models in the candidate set, ΔAIC_{RC} values corresponding to $\alpha = 0.05$ increase. All estimates are based on 1000 simulated phylogenies per *a*.

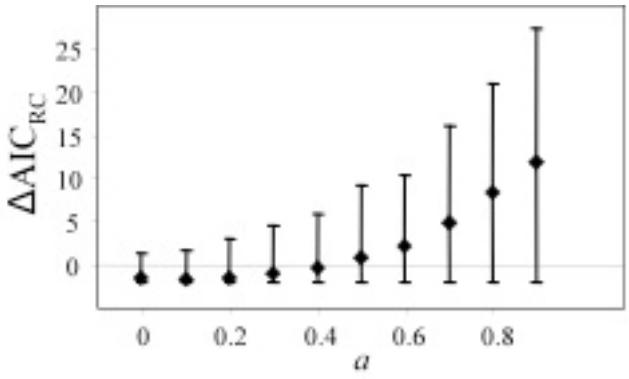

Figure 2.3. Relationship between ΔAIC_{RC} and the extinction fraction a when survival analysis is used to compute likelihoods. Shown are median (black diamonds) and the $0 - 95^{th}$ percentile range of the distribution of ΔAIC_{RC}. Median values and 95^{th} percentile of the ΔAIC_{RC} distribution are positively correlated with a, indicating that this method cannot separate a temporal increase in diversification from a rate-constant model with $a > 0$. All estimates are based on 1000 simulated phylogenies per a.

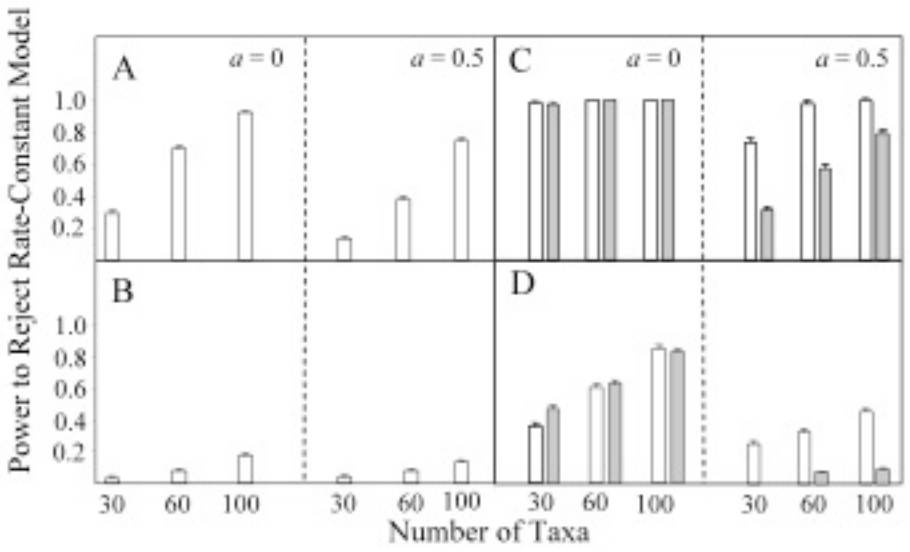

Figure 2.4. Power of BDL (open bars) and the γ-statistic (shaded bars) to detect shifts in diversification rates over time. Test phylogenies were simulated under a model where diversification rates shift with the birth of the N/2 lineage. Power is the percentage of 500 replicate phylogenies for which the null hypothesis of rate-constancy was rejected for each value of *a* and N if the ΔAIC_{RC} rejection region is set to maintain $\alpha = 0.05$ (BDL) or if $\gamma \leq -1.645$. (A) five-fold increase in *r* over time; (B) two-fold increase in *r*; (C) five-fold decrease in *r*; (D) two-fold decrease in *r*. Power for BDL exceeded that of the γ-statistic in all but two of the diversification scenarios considered (4D; $a = 0$; N = 30 and N = 60). Confidence limits are bootstrap standard errors of each power estimate.

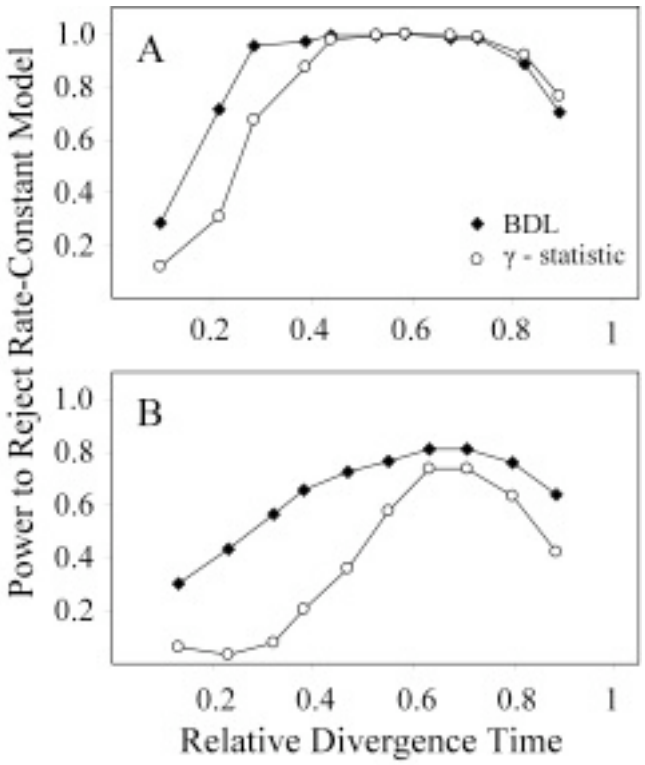

Figure 2.5. Power of BDL and the γ-statistic to detect temporal decreases in diversification rates when rates have shifted at different points in time. Test phylogenies (N = 60) were simulated under (A) $a = 0$ and (B) $a = 0.5$, with a 3.5-fold decrease in the net diversification rate occurring with the birth of the 5th, 10th….50th, or 55th lineage. Power is the percentage of 500 replicate phylogenies for which the null hypothesis of rate-constancy was rejected. Relative divergence times reflect mean the mean time of the rate shift for each set of simulated phylogenies, expressed as a fraction of total clade age. Bootstrap standard errors of each power estimate (not shown) were less than 0.022.

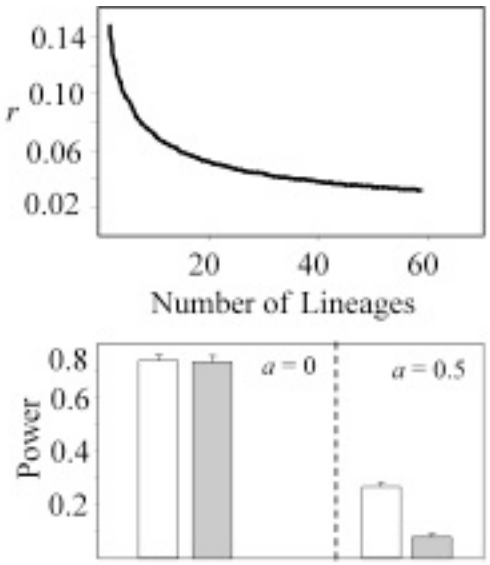

Figure 2.6. Power of BDL (open bars) and the γ-statistic (shaded bars) when diversification rates have decreased gradually under a model of density-dependent cladogenesis. (A) Simulation model: the net diversification rate *r* decreases as the number of surviving lineages increases. (B) Percentage of 500 simulated phylogenies (N = 50) for which the rate-constant model was rejected for likelihood and the γ-statistic under $a = 0$ and $a = 0.5$.

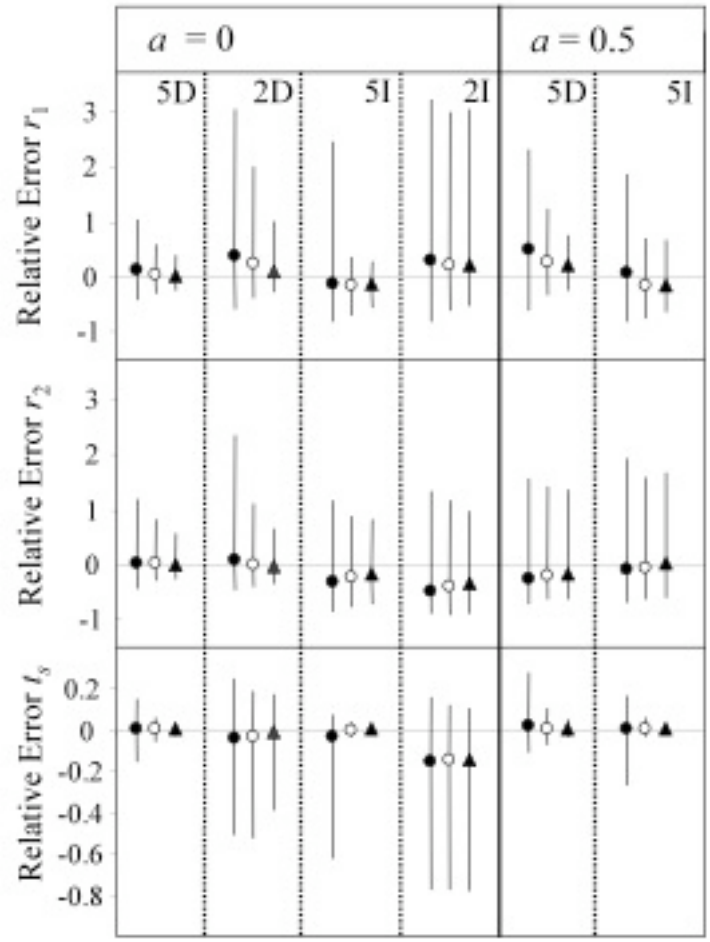

Figure 2.7. Relative error in parameters estimated under the best fit rate-variable model for phylogenies simulated under four models of diversification with $a = 0$ and two models with $a = 0.5$. r_1 and r_2 are maximum likelihood estimates of the net diversification rate before and after the inferred rate shift, t_s. Shown are means and 95% confidence limits on the distribution of relative errors for each set of simulated phylogenies. In general, mean relative errors tended to zero as the number of taxa increased.

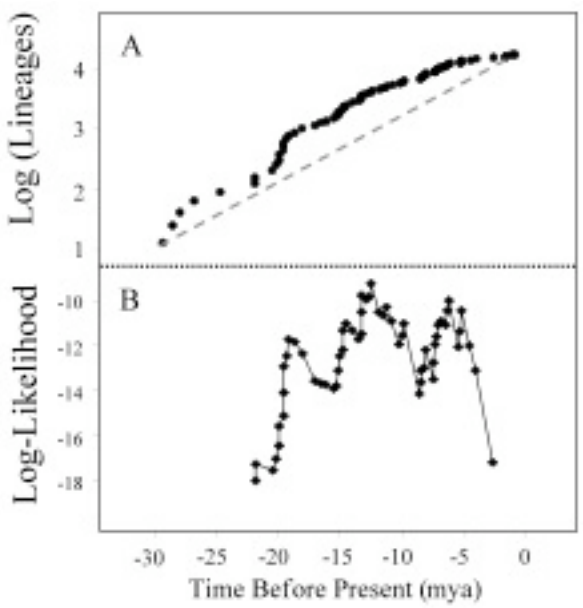

Figure 2.8. Analysis of diversification in Australian agamid lizards. (A) Log-lineage through time plot using the NPRS tree from Harmon et al. (2003). Dashed line represents expected rate of lineage accumulation under the pure birth model. (B) Log-likelihood of rate shifts at different points in time for the agamid data. The best fit model was the three parameter model with two speciation rates and extinction set to zero, with an estimated rate shift occurring 13 million years ago. Estimated speciation rates are 0.148/my and 0.048/my before and after the shift point, respectively, assuming a clade age of 30 million years.

Table 2.1. Results of fitting four birth-death models to the agamid data. ΔAIC is the difference in AIC scores between each model and the overall best-fit model.

Model	*Rate-constant* $a = 0$	*Rate-constant* $a \geq 0$	*Rate-variable* $a = 0$	*Rate-variable* $a \geq 0$
Parameters in model	r	r, a	r_1, t_s, r_2	r_1, t_s, r_2, a
Log-likelihood	-19.34	-19.34	-9.36	-9.36
AIC	40.68	42.68	24.72	26.72
ΔAIC	15.96	17.96	0	2

Table 2.2 — Diversification parameters for the Australian Agamidae under the best-fit model before and after the inferred rate shift (my = million years).

	r	*95% lower*	*95% upper*
Divergences > 13 mya	0.148/my	0.119/my	0.196/my
Divergences ≤ 13 mya	0.048/my	0.039/my	0.063/my

the addition of the five parameter rate variable model, where the ΔAICRC corresponding to α £ 0.05 jumps from 4.8 to 8.2.

As predicted, likelihood methods based on survival analysis show high Type I error rates when the extinction fraction a increases (Figure 2.3). Because available survival models do not incorporate the effects of extinction, they cannot distinguish between temporal increases in diversification rates and constant-rate models with a > 0.

Power Analyses

For each of the 500 phylogenies simulated under a given diversification model, the rate constant model was rejected in favor of a rate-variable scenario if ΔAICRC exceeded that required to maintain a Type I error rate of 0.05 under the pure-birth model (Figure 2.2C). Power to reject the null hypothesis of rate-constancy is generally high when diversification rates have decreased over time (Figure 2.4). Likelihood can separate temporal increases in diversification rates from constant, non-zero extinction, but power to detect this shift is high only when sample sizes are large (N = 60, N = 100) and the magnitude of the rate shift is great. When diversification models included constant background extinction (a = 0.5), power to reject the null hypothesis decreased in all cases.

Under diversification models with extinction set to zero, BDL and the γ-statistic had approximately equal power to detect temporal decreases in diversification rates. When diversification models included non-zero extinction rates (a = 0.5), likelihood methods performed much better than the γ-statistic. BDL showed much greater power to detect temporal decreases in diversification rates when rate shifts occur early in the history of a clade, particularly when extinction is present (Figure 2.5). In the case of density-dependent cladogenesis, power was virtually identical for

BDL and the γ-statistic when a = 0 (Figure 2.6). Likelihood performed better under a = 0.5, but power to reject the null hypothesis was low for both methods. Despite a nearly five-fold decrease in the net diversification rate over time, power was much lower than in the case where rate shifts occurred instantaneously (Figure 2.4).

Parameter Estimation

For models with a = 0, estimates of the net diversification rate showed a weak bias for small phylogenies (Figure 2.7), but the mean relative errors tended to zero as the number of taxa increased. For weak increases or decreases in diversification rates, variance in estimates of r1 and r2 was high. Estimates of the time of the rate shift ts did not show appreciable bias for most of the diversification scenarios considered. For weak increases in diversification rates, mean relative errors in ts suggested a consistent underestimate of true shift times; however, median relative errors in this parameter were less than 0.01 for all three sets of simulated phylogenies (N = 30, 60, 100). For five-fold increases or decreases in diversification rates, error in estimates of ts was very low for phylogenies of N = 60 and N = 100 taxa. For models with a = 0.5, estimates of r1, r2 and ts generally appeared unbiased and tended to zero as the number of taxa increased.

Australian Agamids

The agamid phylogeny was analyzed under two rate-constant and two rate-variable models of diversification (Table 2.1). The best rate-variable model had an AIC of 24.72, versus 40.68 for the best rate-constant model (ΔAICRC = 15.96). The ΔAICRC value required to maintain α = 0.05 was obtained by simulating 1000 phylogenies of the same size as the test phylogeny under the pure-birth model. For N = 69, we reject the null hypothesis of rate-constancy if the observed ΔAICRC ³ 5.0.

ΔAICRC for the agamids was much greater than the ΔAICRC required to reject the null hypothesis at α= 0.05, providing strong support for the hypothesis that diversification rates have changed over time. Diversification rates in the Australian agamids appear to have decreased roughly threefold over time (Table 2.2), with an estimated rate shift occurring 12 – 14 mya. Confidence limits were placed on estimated diversification rates using Moran's variance (Nee 2001).

Log-likelihoods of rate shifts under the best rate-variable model were plotted as a function of time (Figure 2.8). Although the maximum likelihood shift point occurs approximately 13 mya, two additional peaks in the likelihood plot (Figure 2.8B; 6 and 19 mya) suggest that agamid diversification trends may have occurred in a stepped fashion. This pattern of multiple likelihood peaks separated by troughs appears inconsistent with a model of gradually decreasing diversification rates over time, as could occur if the agamid radiation in Australia was characterized by density dependent cladogenesis on a continental scale. Additional work is needed to determine whether this pattern can be attributed to stochasticity or a true stepwise decline in diversification.

Discussion

Model Selection

Type I error rates are high when the model with the lowest AIC score is selected as that which best approximates the data, approaching 50% in some trials. This effect appears to be most pronounced for phylogenies generated under a £ 0.7. Because there is no positive relationship between ΔAICRC and a, we can reject the null hypothesis of rate-constancy if ΔAICRC is greater than or equal to that required to maintain α = 0.05 under the pure-birth model for a given value of N. The dependency of this rejection criterion on the size of the clade (Fig. 2C) and the number

of models under consideration (Fig. 2D) suggests that no single criterion can be used to select amongst among rate-constant and rate-variable models. Researchers studying clades larger or smaller than those considered here may wish to obtain ΔAICRC scores required to maintain an acceptable Type I error rate by simulating phylogenies of the same size as the actual phylogeny under the pure-birth model and examining the resulting distribution of ΔAICRC scores with respect to the full set of candidate models.

Alternative model selection criteria, such as the AICc (Burnham and Anderson 2002) and the likelihood ratio test (LRT) also have high Type I error rates (data not shown), and there seems to be little advantage in abandoning the AIC, provided that error rates and rejection criteria are explicitly addressed through simulation. Even if Monte Carlo methods are used to infer the null distribution of the LRT, this method results in a large number of pairwise LRTs, the significance of which are difficult to compute and interpret (Paradis 1998). When comparing more than two models, one must correct the LRT for multiple testing (Paradis 1998; Pol 2004); this problem is avoided by use of the AIC.

Power Analyses and Parameter Estimation

Despite the high ΔAICRC values required to maintain Type I error rates of 0.05 or less, BDL retains considerable power to detect variation in diversification rates over time. When extinction rates are equal to zero, differences in power for BDL and the γ-statistic are trivial. However, when clades have grown under a model that includes constant, non-zero extinction rates, BDL has much greater power than the γ-statistic.

Decreased diversification through time results in an excess of early diverging lineages in reconstructed phylogenies, but constant background extinction can reduce

this effect by eliminating ancient divergences and favoring recently diverged lineages. By explicitly including extinction, model-based approaches can succeed where other approaches fail.

It is far more difficult to detect a temporal increase in diversification than a temporal decrease. In the birth-death likelihood framework, power to detect weak increases in diversification rates is always low. However, power increases with sample size when the magnitude of the rate shift is large. Furthermore, this may be the only approach available that can detect an increase when it occurs. For most groups, the fossil record does not permit independent estimates of extinction rates, and it is difficult to justify why a particular background extinction rate is less likely than any other; this limits the utility of methods where rejection criteria are dependent on a to cases where diversification rates have decreased over time. Methods that do not explicitly account for extinction, such as the γ-statistic and survival analysis, are unable to distinguish between temporal increases in diversification rates and rate-constant models with $a > 0$.

Under both $a = 0$ and $a = 0.5$, parameter estimation using the birth-death model performs well; estimates of both the timing and magnitude of the rate shift were consistent and generally unbiased across a range of rate-variable diversification scenarios. The two main conclusions from previous work on parameter estimation in a birth-death framework are that (1) it is always difficult to estimate a in the absence of fossils, and (2) we can infer r with reasonable confidence when a is low (Nee et al. 1994a; Paradis 2004). This study extends previous work in two important respects. It is shown that, at least for the diversification models considered here, estimates of the timing of the rate shift are unbiased and have low variance. The sole exception to this pattern occurs when rates show weak increases over time; this can be explained, at least in part, by the low power of BDL to detect such shifts in diversification rates

(Figure 4.4B). This consistency implies that the method will be useful for testing specific hypotheses about the causes of variation in diversification rates. To the extent that we can accurately calibrate divergence times in a molecular phylogeny, we should be able to use the method to examine, for example, whether rate shifts coincide with climatic or geological events that may have influenced the tempo of diversification.

A second point is that the relative error in estimates of the magnitude and timing of rate shifts when diversification rates increase over time perform nearly as well as estimates when rates decrease over time. Confidence limits on the distribution of relative errors under scenarios of increasing and decreasing diversification are similar, particularly for phylogenies of N = 60 and N = 100 taxa. Provided that we can detect a temporal increase in the net diversification rate when it occurs, we can be reasonably confident in the resulting estimates of the magnitude and timing of the rate shift.

Strengths of the birth-death likelihood approach

Primary strengths of the likelihood-based methods discussed here include power to detect both increases and decreases in diversification over time and parameter estimation. Furthermore, the method is very flexible and could accommodate tests of a wide range of rate variable models; although I considered only simple models, net diversification rates could be modeled as functions of any number of independent variables. For example, rates can be modeled as a function of the number of extant lineages at any point in time to test whether decreasing diversification rates show density dependence. Another application might include testing whether diversification rates are related to climatic oscillations (Kadereit et al. 2003). When we have reason to suspect that a particular event in earth history may have affected diversification, rate-variable models can be adjusted to test a priori

hypotheses. Finally, likelihood methods can test for simultaneous rate shifts across multiple, distantly related groups of organisms, provided we can appropriately calibrate divergence times (Turgeon et al. 2005).

Limitations of likelihood methods

Conclusions about the temporal nature of diversification are dependent upon the quality of the data themselves. Here I have ignored error in branching time estimates to focus on the analysis of diversification from these data. In reality, many factors can influence or bias estimates of divergence times. Errors in tree topology per se are not necessarily fatal for likelihood-based analyses of diversification; for example, an ancient, rapid radiation may result in a virtual star topology, with all lineages appearing to originate simultaneously. It may never be possible to reconstruct the true phylogeny for these taxa, but it may nonetheless be possible to infer with confidence that diversification rates have decreased over time.

Biased estimates of divergence times, on the other hand, are expected to introduce a similar bias into diversification rate analyses. For example, failure to use an appropriate model of sequence evolution to estimate branch lengths could lead to consistent underestimates of divergence times early in the history of a clade. This is expected to occur when simple models of sequence evolution cannot fully capture the effects of saturation that become apparent when comparing highly diverged lineages (Arbogast et al. 2002). In this case, one might conclude that decreasing diversification rates have prevailed, when the pattern is a simple artifact of saturation (Revell et al. 2006).

Incomplete taxon sampling is a potentially severe problem that has not been addressed for BDL. If a number of taxa in a clade are not sampled, and if sampled lineages are random with respect to divergence times, we will observe a spurious

decrease in diversification over time, because missing branching times tend to occur closer to the present than the root of the tree (Nee et al. 1994a). Monte Carlo approaches can be used to generate the null distribution of branching times with incomplete sampling, if the true number of species in the clade is known (Pybus and Harvey 2000; Pybus et al. 2002). A similar approach could be used with BDL to obtain the ΔAICRC score required to reject a rate-constant model while minimizing Type I error rates.

The BDL approach could be adapted to test a variety of birth-death models other than the simple rate shift models considered here. In the case of density-dependent cladogenesis (Figure 4.4), one might predict that power to reject the null hypothesis would increase if a density-dependent model had been included in the candidate set. However, Type I error rates increase with the number of rate-variable models considered (Figure 2.2D); a greater difference in AIC scores between the best rate-constant and rate-variable models is required to reject the null hypothesis when the number of fitted models is increased. Because Type I error rates do not increase with a, we can easily determine appropriate rejection criteria for the null hypothesis under any set of candidate models by simulating rate-constant phylogenies and examining the distribution of ΔAICRC values.

The relationship between hypothesis testing and data exploration can easily become blurred when applying these methods, and this is a particular concern when using a priori hypotheses to test models of diversification. If, before looking at any data, we have reason to believe that diversification rates may have shifted at some point in time, we can justifiably reduce the number of free parameters in rate-variable models by fixing the timing of the rate shift to the hypothesized value (Near et al. 2003; Turgeon et al. 2005). This reduces the AIC penalty term of the rate-variable model relative to the rate-constant model, and we have greater power to reject the null

hypothesis of rate-constancy if it is false. However, there is danger that researchers will simply look at a phylogenetic tree or a plot of lineages-through-time, note an apparent increase or decrease in diversification, and then construct an a posteriori hypothesis to explain this pattern. If this a posteriori observation is then treated as an a priori hypothesis in a model-based analysis, the probability of a Type I error may be increased. This is a subtle but important point, and researchers must be careful to justify why a particular a priori hypothesis was used.

Summary and Recommendations

The analyses presented here indicate that the birth-death likelihood (BDL) approach is a useful framework for studying temporal variation in diversification rates. In all scenarios considered, BDL showed comparable or greater power than the γ-statistic to detect temporal variation in diversification rates. The advantages afforded by BDL are particularly apparent when extinction is present, suggesting that researchers may wish to revisit datasets for which the γ-statistic was unable to detect a temporal shift in the diversification rate. Furthermore, BDL appears to be unique among available methods in that it can separate a temporal increase in diversification from constant, non-zero extinction rates. Finally, BDL provides parameter estimates that can be used to formulate more specific hypotheses about underlying processes that have influenced the tempo of diversification. However, model overfitting is a potentially serious problem, and researchers must carefully address this issue before concluding that the tempo of diversification has changed over time.

CHAPTER 3

LASER: A MAXIMUM LIKELIHOOD TOOLKIT FOR DETECTING TEMPORAL SHIFTS IN DIVERSIFICATION RATES FROM MOLECULAR PHYLOGENIES

Abstract

Rates of species origination and extinction can vary over time during evolutionary radiations, and it is possible to reconstruct the history of diversification using molecular phylogenies of extant taxa only. Maximum likelihood methods provide a useful framework for inferring temporal variation in diversification rates. LASER is a package for the R programming environment that implements maximum likelihood methods based on the birth-death process to test whether diversification rates have changed over time. LASER contrasts the likelihood of phylogenetic data under models where diversification rates have changed over time to alternative models where rates have remained constant over time. Major strengths of the package include the ability to detect temporal increases in diversification rates and the inference of diversification parameters under multiple rate-variable models of diversification. The program and associated documentation are freely available from the R package archive at http://cran.r-project.org.

Introduction

Recent years have seen an explosive proliferation of DNA sequence data for molecular phylogenetic analyses and a commensurate increase in the use of these data to draw inferences about macroevolutionary processes. A particularly active area of research involves the use of molecular phylogenies to study variation in rates of species origination and extinction, both among lineages (Slowinski and Guyer 1989; Mooers and Heard 1997) and over time (Nee et al 1994; Paradis 1997).

Likelihood Analysis of Speciation and Extinction Rates (LASER) is a package for the R programming environment that facilitates model-based analyses of diversification rates. LASER is the first software package to implement tests for temporal variation in diversification rates using likelihood methods based on the birth-death process (Nee et al 1994). LASER is licensed under the GNU General Public License and complements the existing R libraries 'ape' (Paradis et al 2004) and 'apTreeshape' (Bortolussi et al 2006), which provide functions for phylogenetic tree manipulation and the analysis of among-lineage heterogeneity in diversification rates.

LASER was written to address several limitations of existing software for analyzing the tempo of diversification. Approaches such as the gamma statistic (Pybus and Harvey 2000) and survival analysis (Paradis 1997), which are implemented in the R library 'ape' (Paradis et al 2004), test for departures from the pure-birth model of cladogenesis, and can only be used to infer temporal decreases in diversification rates (Nee 2001; Rabosky 2006). These methods are thus unable to address many questions of interest to evolutionary biologists, such as whether temperate faunas experienced elevated speciation rates during the Pleistocene (Weir and Schluter 2004). Furthermore, existing methods suffer reduced power to detect temporal decreases in diversification rates when clades have diversified under high background extinction rates (Rabosky 2006).

LASER fits a candidate set of rate-variable diversification models to phylogenetic data and contrasts the likelihood of the data under these models to alternatives where speciation and extinction rates have remained constant over time. The null hypothesis that diversification rates have not changed over time is tested using the statistical approach described in Rabosky (2006). The test statistic for constancy of diversification rates is computed as

(3.1) $\Delta AIC_{RC} = AIC_{RC} - AIC_{RV}$

where AIC_{RC} is the Akaike Information Criterion (AIC) score for the best-fit rate-constant model of diversification, and AIC_{RV} is the AIC score for the best-fit rate-variable model under consideration. Thus, a positive ΔAIC_{RC} value suggests that the data are best approximated by a rate-variable model of diversification. Although several previous studies have used the AIC to distinguish among rate-constant and rate-variable models of diversification (Barraclough and Vogler 2002; Turgeon et al 2005), Rabosky (2006) found that this approach results in high Type I error rates unless critical values of the ΔAIC_{RC} distribution are explicitly addressed through simulation.

The LASER package provides a comprehensive toolkit for computing ΔAIC_{RC} for test phylogenies and for comparing the observed ΔAIC_{RC} statistic to its distribution under the null hypothesis. This is the first available approach that can detect temporal increases in diversification rates, and extensive simulation has shown that the method has greater power than other methods to detect temporal declines in diversification rates when clades have diversified under elevated background extinction rates (Rabosky 2006).

Additional strengths of the model-fitting approach implemented in the LASER include the ability to test hypotheses of rate variation while estimating relevant diversification parameters. Furthermore, the package can be used to test *a priori* hypotheses of temporal rate variation. The R programming environment used by LASER provides great flexibility, and likelihood functions in the package can easily be tailored to a variety of statistical applications. For example, one can generate posterior distributions of diversification parameters using the posterior distribution of

phylogenetic tree topologies and branch lengths sampled using Markov chain Monte Carlo (MCMC) methods (eg Huelsenbeck and Ronquist 2001).

Usage

LASER operates on sets of branching times derived from ultrametric phylogenetic trees, and provides functions for obtaining branching times from several input formats, including the widely used 'Newick' (parenthetic) tree format. Likelihoods and parameter estimates can be obtained for a range of rate-variable diversification models, including logistic and exponential density-dependent models and multi-rate birth-death models. Additional functions permit batch processing of multiple phylogenies to obtain the null distribution of ΔAIC_{RC} or posterior distributions of diversification parameters.

The function fitdAICrc computes the ΔAIC_{RC} test statistic for a test phylogeny using arguments that specify the candidate set of rate-variable models to be considered. The null distribution of the test statistic is obtained by either simulating branching times with the function yuleSim or by importing simulated trees using the function getBtimes.batch. The latter function is particularly useful for the analysis of phylogenies with incomplete taxon sampling, because incomplete sampling can result in a spurious decline in diversification rates over time (Pybus and Harvey 2000). To address this problem in LASER, one can simply generate rate-constant phylogenies with incomplete sampling using PhyloGen (Rambaut 2002) or other software and import the trees into LASER to tabulate the null distribution of the ΔAIC_{RC} test statistic.

A call to the function fitdAICrc.batch will then generate the null distribution of the test statistic and return the probability of the observed ΔAIC_{RC} index under the null hypothesis. Functions are available to call any rate-variable and rate-constant

diversification models individually, and additional functions permit exploration of diversification patterns for any user-defined temporal interval.

Diversification Models

Rate-constant diversification models implemented in LASER include the pure birth model, with a constant speciation rate $\lambda > 0$, and the birth-death model, with $\lambda > 0$ and extinction rate $\mu \geq 0$. Seven rate-variable diversification models are provided, including density-dependent and multi-rate variants of the pure birth and birth-death models. The package includes both logistic and exponential density-dependent speciation models. Under the logistic density-dependent model of cladogenesis, the speciation rate λ at time t is modeled as

$$\lambda(t) = \lambda_o\left(1 - \frac{N_t}{K}\right) \tag{3.2}$$

where λ_0 is the initial speciation rate, N_t is the number of lineages at time t in a reconstructed phylogeny, and K is analogous to the carrying capacity parameter of population ecology. Speciation rates are modeled under an exponential density-dependent process as

$$\lambda(t) = \lambda_o N_t^{-x} \tag{3.3}$$

where x controls the magnitude of the rate change with respect to the number of lineages at any point in time in the reconstructed phylogenetic tree.

Multi-rate variants of the pure birth and birth-death model assume the existence of one or more breakpoints in time, such that a clade has diversified under one set of diversification parameters before the breakpoint and another set of

parameters after the breakpoint. For example, LASER includes a two-rate pure birth model with three parameters: the initial speciation rate, the final speciation rate, and the time of the rate shift.

Example: Holarctic Damselfly Radiation

Turgeon et al (2005) tested whether Holarctic damselflies in the genus *Enallagma* showed evidence for increased diversification rates during the Quaternary. They found evidence for a recent increase in speciation rates, suggesting a role for Pleistocene glacial cycles in damselfly diversification. Their conclusions were based on a model-fitting approach similar to that described above. However, as noted in Rabosky (2006), this method can result in high Type I error rates. To explicitly address this problem, I used LASER to compute ΔAIC_{RC} for the *Enallagma* phylogeny from Turgeon et al (2005) and to tabulate the null distribution of the statistic as follows:

```
data.bt <- getBtimes(file = 'enallagma.tre')
summary <- fitdAICrc(data.bt)
```

The first line creates a vector data.bt of the branching times for *Enallagma* by reading the parenthetic tree stored in enallagma.tre. The function fitdAICrc generates an object, summary, that contains the results of fitting all rate-variable and rate-constant models to the data. The observed ΔAIC_{RC} statistic for *Enallagma* is 13.0372, and the best fit model is a three-parameter rate-variable model specifying a 12.5-fold increase in the speciation rate over time. The significance of the observed ΔAIC_{RC} statistic was assessed with the following commands:

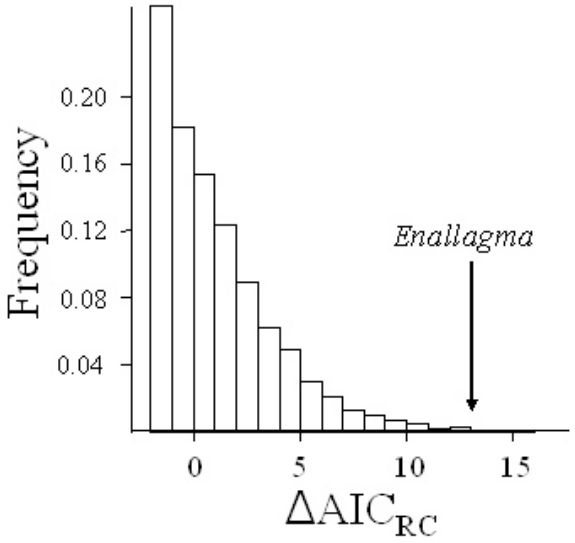

Figure 3.1. Distribution of the ΔAIC_{RC} test statistic for 5000 rate-constant phylogenies of the same size as the *Enallagma* phylogeny. The calculated ΔAIC_{RC} for *Enallagma* was 13.0372 and indicates a highly significant temporal increase in the net diversification rate over time ($p = 0.0016$).

```
null.bt <- yuleSim(37, 5000)

fitdAICrc.batch(null.bt, stat = 13.0372)
```

The first line simulates 5000 phylogenies with the same number of tips as the *Enallagma* tree (37) under the null hypothesis of rate-constancy and stores them in object null.bt. fitdAICrc.batch analyzes the set of simulated phylogenies and approximates the probability of the observed ΔAIC_{RC} statistic under the null hypothesis. In this case, the observed ΔAIC_{RC} statistic indicates a highly significant departure from the null hypothesis of rate-constancy ($p = 0.0016$; Fig. 3.1), supporting the conclusions of Turgeon et al (2005).

Summary

LASER fits multiple rate-variable and rate-constant models of diversification to reconstructed phylogenies using maximum likelihood. Its main strength includes the use of Monte Carlo simulation to control for elevated Type I error rates associated with likelihood-based analyses of diversification. LASER is the first available package that can detect temporal increases in diversification rates, and has considerable power to detect temporal declines in diversification rates when clades have diversified under high background extinction rates. As a freely available package for the R programming environment, it is flexible and platform-independent, and can easily be tailored to a variety of user-specific applications.

Since the original LASER package was described, a number of additional functionalities have been added. Version 1.0 contained only the birth-death likelihood methods described in Rabosky (2006). However, LASER 2.0 now includes the following: (1) Implementation of the Monte Carlo Constant Rates (MCCR) test, as

described in Pybus and Harvey (2000); (2) The SPVAR, BOTHVAR, and EXVAR models from Rabosky and Lovette (2008b); (3) the 1-rate and 2-rate rate shift models as described in Rabosky et al. (2007b). The MCCR test implemented in LASER can perform simulations assuming incomplete taxon sampling. The continuous-time models (SPVAR and related functions) model continuous time variation in speciation and extinction rates and represent, as of this writing, the only implementation available that can perform these analyses.

CHAPTER 4

EXPLOSIVE EVOLUTIONARY RADIATIONS: DECREASING SPECIATION OR INCREASING EXTINCTION THROUGH TIME?

Abstract

A common pattern in time-calibrated molecular phylogenies is a signal of rapid diversification early in the history of a radiation. Because the net rate of diversification is the difference between speciation and extinction rates, such 'explosive-early' diversification could result either from temporally declining speciation rates or from increasing extinction rates through time. Distinguishing between these alternatives is challenging but important, because these processes likely result from different ecological drivers of diversification. Here we develop a method for estimating speciation and extinction rates that vary continuously through time. By applying this approach to real phylogenies with explosive-early diversification and by modeling features of lineage-accumulation curves under both declining speciation and increasing extinction scenarios, we show that a signal of explosive-early diversification in phylogenies of extant taxa cannot result from increasing extinction and can only be explained by temporally declining speciation rates. Moreover, whenever extinction rates are high, 'explosive early' patterns become unobservable, because high extinction quickly erases the signature of even large declines in speciation rates. Although extinction may obscure patterns of evolutionary diversification, these results show that decreasing speciation is often distinguishable from increasing extinction in the numerous molecular phylogenies of radiations that retain a preponderance of early lineages.

Introduction

A central question in evolutionary biology concerns the extent to which species-level diversification rates vary among lineages and over time. This issue has a venerable history in the paleontological literature (e.g., Simpson 1953; Raup 1985). More recently, the increasing availability of robust molecular phylogenies for clades of extant species has generated a surge of interest in methods to extract information about the tempo and mode of evolutionary diversification from them (Nee et al. 1994a; Paradis 1997; Nee 2006; Rabosky et al. 2007).

Because these statistical tools permit inferences about temporal variation in species-level diversification rates, many studies have applied them in association with time-calibrated phylogenies of extant taxa to characterize rates of lineage accumulation through time during evolutionary radiations. One of the most commonly observed patterns in these studies of diversification rates in extant clades is evidence for bursts of diversification in the early stages of those species-level radiations, followed by declining diversification through time. Such 'explosive-early' radiations have been reported from a wide range of taxa and biogeographic settings (e.g., Lovette and Bermingham 1999; Harmon et al. 2003; Shaw et al. 2004; Kozak et al. 2006).

Several alternative ecological hypotheses might explain a pattern of explosive-early diversification reconstructed from a phylogeny of extant species. For example, opportunities for speciation during adaptive radiation might be inversely related to the number of potentially competing species in existence at any point in time; this model of resource-limited diversification would predict that speciation rates should decline in a density-dependent fashion (Walker and Valentine 1984; Nee et al. 1992; Phillimore and Price 2008). Other models implicitly suggest that extinction rates might increase during the course of evolutionary radiations (e.g., Ricklefs and Cox 1972; Levinton 1979; Hubbell 2000). But because the net rate of diversification is simply the

difference between speciation and extinction rates, an increase in the extinction rate could in principle result in precisely the same net diversification rate through time as a decline in the speciation rate. Weir (2006) used a simulation study to suggest that declining speciation was more likely to explain temporal decreases in diversification rates in neotropical avifaunas, but the generality of this result and underlying mechanisms remain untested.

Here we explore whether evolutionary radiations characterized by explosive-early diversification are more likely to have resulted from declining speciation rates or from increasing extinction rates through time. We develop an analytical framework based on the birth-death process (Kendall 1948; Nee et al. 1994b) that explicitly models speciation and extinction rates that vary continuously through time. We apply these methods to three published phylogenies that have in common a strong pattern of lineage accumulation consistent with early, rapid diversification, and we test whether models specifying constant speciation and time-varying extinction provide a better fit to real data than models of time-varying speciation and constant extinction. We further use simulations of declining diversification through time to contrast features of lineage accumulation curves under scenarios of decreasing speciation and increasing extinction through time.

Methods

Modeling framework

To test whether temporal declines in diversification rates are best explained by changes in speciation or extinction rates, we require a modeling framework for speciation and extinction rates that vary continuously through time. Consider a general birth-death process, where existing lineages give birth to new lineages at a per-lineage rate λ and go extinct with rate μ. The general probability model described below was

developed by Nee et al. (1994b); while this framework has not yet been used for inference on time-varying speciation and extinction rates, there is nothing in this model that prohibits λ and μ from varying over time or among lineages.

A simple way to model the growth of a phylogenetic tree through time is to 'split' the tree into a collection of daughter branches, with each branch originating at some time t_i and surviving to the present day (time T). Here we consider only the reconstructed evolutionary process (Nee et al. 1994b), where all lineages survive to the present; this corresponds to a typical molecular phylogeny, because only those species that have not gone extinct are observed in a phylogeny that includes only extant taxa. Let $\lambda(t)$ and $\mu(t)$ denote time-varying speciation and extinction rates. We are concerned here with temporal variation in lineage diversification rates; although this model can be extended to include among-lineage rate variation, in the model below $\lambda(t)$ and $\mu(t)$ are constant among lineages that exist at time t.

It is convenient to partition the stochastic processes contributing to the likelihood of a phylogenetic tree with N taxa under $\lambda(t)$ and $\mu(t)$ into two components. The first is attributable to speciation events: new lineages arise in a growing clade with a probability proportional to

(4.1) $(i - 1)\ \lambda(t)\ \mathrm{P}(t_i, T)$

where $\mathrm{P}(t_i, T)$ is the probability that a lineage in existence at time t_i will survive to be observed at time T (e.g, the lineage will not go extinct). The $(i - 1)$ term comes from the fact that, immediately prior to the birth of the i'th lineage at time t_i, the tree contains a total of $(i - 1)$ lineages that could potentially give birth. The second component of the likelihood follows from the observation that each of N lineages survives from some time t_i to T, leaving only a single descendent in the present

(itself). It may seem counterintuitive to imagine each lineage leaving only a single progeny lineage, but we are modeling the growth of the phylogenetic tree as a collection of such processes (Figure 4.1).

Define P(t_i, T), or the probability that a lineage survives between time t_i and T, as:

$$(4.2) \quad P(t_i,T) = \left[1 + \int_{t_i}^{T} \mu(\tau)\exp\left(\rho(\tau,t_i)\right)d\tau\right]^{-1}$$

(Kendall 1948), where

$$(4.3) \quad \rho(\tau,t_i) = \int_{t_i}^{\tau} \left\{u(s) - \lambda(s)\right\}ds$$

When μ and λ are constant through time, μ / λ represents the long-term probability that a lineage goes extinct (e.g., Raup 1985), and it is true that

$$(4.4) \quad \lim_{T\to\infty} P(t_i,T) = \mu/\lambda$$

To calculate the probability that each lineage i leaves a single progeny (itself) on the interval (t_i, T), we note that the number of progeny lineages under the birth-death process follows a geometric distribution. The probability that a birth-death process beginning with a single lineage will result in k surviving lineages after some time T is given by $(1 - u)^k$, where 1 - u is the parameter of the geometric distribution of progeny lineages (Nee et al. 1994b). We denote the probability that lineage i leaves a single progeny in the present as ξ_i , where

$$(4.5) \quad \xi_i = P(t_i,T)\exp\left[\rho(T,t_i)\right]$$

Combining equations 4.1 and 4.5, we obtain the likelihood for N lineages:

$$(4.6) \quad L = (N-1)! \prod_{i=3}^{N} \{\lambda(t_i) P(t_i, T)\} \prod_{i=3}^{N} \{\xi_i\} \{\xi_2^2\}$$

which is identical to Nee et al. (1994b; eqn 20). Note that the t_i's are simply the speciation times (Figure 4.1). Equation 6 considers only $N - 2$ speciation events, because the first two speciation events must have occurred; if they had not, no phylogenetic tree would exist to be observed (Nee et al. 1994b). The ξ_2 term in eqn 4.6 corresponds to these two basal branches. The likelihood function (eqn 4.6) frequently results in positive log-likelihood values; this occurs because the $(i - 1)$ $\lambda(t)$ $P(t_i, T)$ component of the likelihood is not normalized and is merely proportional to the actual probability density.

Models for declining diversification rates

Our general approach is to ask whether the pattern of lineage accumulation through time in a molecular phylogeny is best explained by a model with constant speciation and time-varying extinction, or by a model with constant extinction and time-varying speciation. The first step of the process is to choose appropriate models for $\lambda(t)$ and $\mu(t)$. We used a simple exponential model, under which the time varying speciation rate is given by

$$(4.7) \quad \lambda(t) = \lambda_0 \exp(-kt)$$

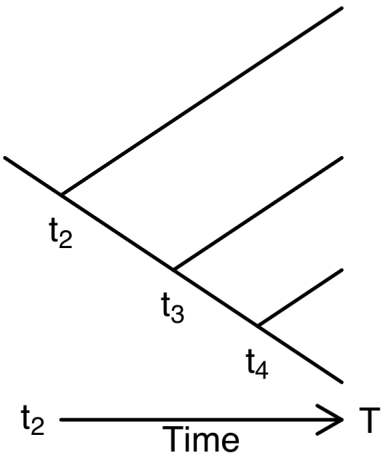

Figure 4.1. Reconstructed phylogenetic tree illustrating parameters described in text. Time is measured from the root to the present (time T), with each t_i corresponding to the speciation time of the i'th lineage. By definition, a clade originates with the birth of a second lineage at time t_2; note that the two basal lineages that define the clade persist from time t_2 to T.

where λ_0 is the initial speciation rate and k specifies the magnitude of the rate decline through time ($0 \leq k < \infty$). When k = 0, the speciation rate is constant through time. The time-varying extinction rate was modeled as

$$\mu(t) = \mu_0\left(1 - \exp\left[-zt\right]\right) \tag{4.8}$$

where μ_0 is the asymptote of the increasing extinction rate through time and z controls the steepness of the increase in extinction with respect to time ($0 < z < \infty$). When z is very large, the extinction rate is a constant μ_0 through time. These models are flexible and can accommodate a range of declining diversification scenarios.

We assumed that declining net diversification rates through time could result from three general processes (i) declining speciation through time but constant extinction (a three parameter model: λ_0, k, and μ); (ii) increasing extinction through time, but constant speciation (three parameters: μ_0, z, and λ); and (iii) declining speciation rates and increasing extinction rates through time (four parameters: λ_0, k, μ_0, and z). For clarity, we refer to these models as SPVAR (time-varying speciation only), EXVAR (time-varying extinction only), and BOTHVAR (both speciation and extinction vary through time). Thus, for the SPVAR model, the net diversification rate $r(t)$ is given by

$$r(t) = \lambda_0 \exp(-kt) - \mu_0 \tag{4.9}$$

and for the EXVAR model,

$$r(t) = \lambda_0 - \mu_0\left(1 - \exp\left[-zt\right]\right) \tag{4.10}$$

and for BOTHVAR,

$$r(t) = \lambda_0 \exp(-kt) - \mu_0\left(1 - \exp\left[-zt\right]\right) \quad (4.11)$$

We constructed likelihood functions for SPVAR, EXVAR, and BOTHVAR models by finding analytical solutions to eqn (4.3) and substituting the relevant expression into eqns (4.2) and (4.6). For nonlinear models of speciation and extinction, there is generally no analytical solution to the integral in eqn (4.2); we performed the required numerical integrations using the QUADPACK-derived routine (Piessens et al. 1983) as implemented in the function 'integrate' for the R programming environment (http://cran.r-project.org/). Models were fitted to phylogenetic data using a box-constrained derivation of the BFGS quasi-Newton method (Byrd et al. 1995). This enabled us to enforce constraints on parameters to meet assumptions of the model, specifically the fact that extinction rates cannot exceed speciation rates. Optimization was performed in R using the function 'optim' with the 'L-BFGS-U' method. Because optimization of the likelihood function can fail when multiple optima are present, we repeated all optimization procedures 100 times with random starting parameter values. All source code for numerical fitting of SPVAR, EXVAR, and BOTHVAR models has been placed in the R package LASER (Rabosky 2006a).

Application to data

To determine whether patterns of diversification during explosive-early radiations are best explained by changes in speciation or extinction rates, we applied the method to three published phylogenies: (i) Australian lizards in the family Agamidae (Harmon et al. 2003); (ii) North American wood-warblers in the genus

Dendroica (Lovette and Bermingham 1999); and (iii) Australo-Papuan pythons (Rawlings et al. 2008). These three radiations all show a phylogenetic pattern of explosive-early diversification, followed by declining diversification rates through time. We selected these studies because the phylogenetic trees used in each case are 93+% complete at the species level, reducing the risk of detecting spurious declines in diversification rates due to incomplete taxon sampling (Pybus and Harvey 2000), and because conclusions about declining diversification rates were previously inferred in each case by at least two different methods (e.g., Pybus and Harvey 2000; Rabosky 2006b).

For each group, we obtained the ultrametric trees used to produce the lineage-through-time (LTT) plots that appeared in the original papers. All data were rescaled such that the basal divergence occurred 1.0 time units before the present, and we then fitted the three rate-variable diversification models (SPVAR, EXVAR, BOTHVAR) to each tree. For comparison with the constant-rate diversification process, we also fitted each tree with a simple two-parameter birth death model, where $\lambda(t) = \lambda$ and $\mu(t) = \mu$. We could not use the likelihood ratio test to compare models because the SPVAR and EXVAR models are not nested; rather, we compared model fits using the Akaike Information Criterion (AIC).

Qualitative features of lineage accumulation curves

We also employed simulations to investigate features of lineage accumulation curves when net diversification rates decline through time. We simulated phylogenetic trees under a model of temporally decreasing diversification, where the decline was caused by either decreasing speciation rates or increasing extinction rates. Most previous studies that have simulated time-varying diversification processes have used discrete-time phylogenetic simulation algorithms (e.g., Paradis 1997; Rabosky 2006b),

in which phylogenetic trees are generated by iterating over a series of time steps such that each lineage has a probability of giving birth or going extinct each time step. Because the discrete-time approach is merely an approximation of the continuous-time diversification process, we implemented a simulation procedure that enables phylogenies to be simulated in continuous time with time-varying diversification parameters.

For a given diversification model (e.g., SPVAR) and magnitude of rate change (e.g., a 10-fold reduction in the net diversification rate through time), we found parameters that would – on average - result in a target number of lineages after t = 1.0 time units. We then divided the total simulation time into 50 intervals of t = 0.02 time units and calculated mean values of λ and μ for each interval given the overall diversification parameters λ_0, k, μ_0, and z. Each simulation was initiated with two lineages, which had parameters λ_1 and μ_1 on the first time interval; after t = 0.02 time units, parameters were updated to λ_2 and μ_2 and the simulation was continued to the end of the second time interval (overall elapsed time of 0.04 time units). These sequential parameter updates were continued until the end of the simulation. Thus, while we used a discrete approximation to model and track variation in λ and μ, the underlying simulation occurred in continuous time. All phylogenetic simulation was conducted using a modified version of the birth-death tree simulation algorithm from the Geiger package for R (Harmon et al. 2007).

We simulated phylogenies undergoing 5-fold and 15-fold declines in net diversification rates through time assuming the following diversification models: (i) declining speciation through time, but zero extinction; (ii) declining speciation through time, with high (constant) background extinction; and (iii) increasing extinction through time, with constant speciation. For each scenario, we found λ_0, k, μ_0, and z parameters that would result in an expected number of 80 lineages per simulation

using equations 7-11 (Table 4.1). We then performed 1000 simulations under each diversification model; to reduce any potentially confounding effects of very small or very large phylogenies (e.g., Price 2008), we retained only those simulations that contained between 40 and 120 surviving lineages at the end of the simulation. Parameters used for each diversification scenario are given in Table 4.1, and diversification curves illustrating temporal changes in λ, μ, and r under the simulation model are shown in Figure 4.2.

To test the extent to which temporal declines in net diversification rates can be inferred from phylogenies generated under time-varying speciation and extinction models, we computed the γ-statistic (Pybus and Harvey 2000) for each batch of simulated phylogenies. This statistic provides a convenient summary of the distribution of internode distances in a phylogenetic tree; under a constant rate diversification process with $\mu = 0$, γ follows a standard normal distribution. A constant rate diversification process with $\mu > 0$ will result in $\gamma > 0$. However, only temporal declines in diversification rates can result in $\gamma < 0$ (Pybus and Harvey 2000).

Results

Reconstructed speciation and extinction rates in real phylogenies

The three model phylogenies show pronounced evidence for temporally declining diversification rates. Calculated γ-statistics for each phylogeny are significantly less than zero, and thereby strongly reject both constant-rate diversification processes and temporally increasing diversification rates (agamids: $\gamma = -4.50$, $p < 0.001$; warblers: $\gamma = -4.20$, $p < 0.001$; pythons: $\gamma = -3.15$, $p < 0.001$). In each case, the SPVAR model fit the observed pattern of speciation much better than both the constant-rate birth-death model and the EXVAR model (Table 4.2). More surprisingly, even the constant rate birth-death model consistently fit the data better

than the EXVAR model. Despite a pronounced decline in the speciation rate inferred under the SPVAR model (Table 4.2; $\Delta\lambda$), the change in extinction through time under the EXVAR model (Table 4.2; $\Delta\mu$) was zero for all three datasets, indicating that the best-fit parameterization of this model does not differ from a constant-rate birth-death model. This result shows that the 'explosive early' pattern seen in these topologies cannot be explained by an increase in extinction rates in the more recent period of the radiations.

Likelihoods under the four parameter BOTHVAR model were identical to those under the three parameter SPVAR model (Table 4.2). This is possible because the SPVAR model is simply a special case of the BOTHVAR model with constant extinction through time; thus, if no change in extinction is inferred under BOTHVAR, likelihoods should be identical to those under SPVAR. Reconstructed speciation and extinction through time curves under the BOTHVAR model suggest that speciation rates in agamids, warblers, and pythons have decreased markedly during the course of these radiations, with rates in warblers undergoing the most severe decline (Figure 4.3). These curves are virtually indistinguishable from those inferred under the SPVAR model and specify extinction rates that are at most only marginally greater than 0. The only (minor) exception occurs in the warblers, where we found a trivial increase in the extinction rate through time ($\Delta\mu = 0.23$, but compare with $\Delta\lambda = -8.87$); for the warblers as for the other clades, the BOTHVAR model provides poorer fit than does the SPVAR model (ΔAIC = 2.0; Table 4.2). Because SPVAR and BOTHVAR differ by only a single parameter, and because SPVAR is a special case of BOTHVAR, it is not possible to obtain ΔAIC in favor of SPVAR greater than the observed value of 2.0. These patterns suggest that the changes in net diversification rates through time in these groups have been mediated almost entirely by declining speciation rates and not by increasing extinction rates.

Speciation and extinction rate simulations

Our simulations show that patterns of lineage accumulation through time during explosive-early radiations vary dramatically depending on whether declining diversification rates are a function of decreasing speciation or increasing extinction rates (Figure 4.4). When speciation rates decrease through time, the number of surviving lineages in existence at any point in time is greater than the expected number of lineages under a constant-rate diversification process (Figure 4.4A, B). However, this excess of lineages is replaced by a sigmoidal relationship that much more closely mimics the null pattern when comparable changes in the net diversification rate are driven by increasing extinction rates (Figure 4.4C, D). Under both high extinction scenarios, a modest excess in the number of lineages during the earliest stages of a radiation switches to a modest paucity of lineages later in the radiation, where the lineage-through-time curve for more recent divergences shows the upturn thought to be characteristic of increasing diversification rates through time or high relative extinction rates (Nee et al. 1994a; Rabosky 2006b). This sigmoidal pattern in the simulated LTT plots is especially striking when large increases in net diversification rates are driven solely by increasing extinction rates through time (e.g., Figure 4.4D).

When phylogenies are simulated under a model of declining speciation rates with no extinction, the γ-statistic gives the expected result: larger declines in speciation rates result in lower γ values (Figure 4.5A, B). However, when speciation rates decline under high but constant extinction, the signature of explosive-early diversification is absent (Figure 4.5C, D). For a modest 5-fold decrease in the speciation rate, γ is significantly greater than zero when extinction is relatively high ($t = 6.155$; $df = 999$; $p < 0.001$). Under a 15-fold decline in the net diversification rate

with high but constant extinction, the majority of the distribution of γ lies within the 95% confidence interval for a constant rate diversification process with no extinction (Figure 4.5D). No signature of declining net diversification rates can be detected with the γ-statistic when the rate decrease is attributable to increasing extinction through time (Figure 4.5E, F); this is a particularly striking pattern when compared to identical changes in net diversification rates attributable to declining speciation only (Figure 4.5A, B).

Discussion

We developed and explored a framework for modeling time-varying speciation and extinction rates and for testing whether the pattern of explosive-early diversification seen in many evolutionary radiations is best explained by declining speciation rates or by increasing extinction rates. Although these competing models can result in identical net diversification rates through time, our results indicate that only declining speciation rates leave a signature of rapid lineage accumulation early in the history of radiations that can be inferred from molecular phylogenies that include only extant taxa. We analyzed three representative phylogenies known to show this pattern of rapid, early lineage accumulation and found that a model specifying temporally-declining speciation rates provided a much better fit than a model with increasing extinction rates. More surprisingly, the model specifying increasing extinction rates through time failed to fit the data better than a simple constant rate birth-death model (Table 4.2). In each case, maximum likelihood parameter estimates under the variable extinction model specified no change in the extinction rate through time.

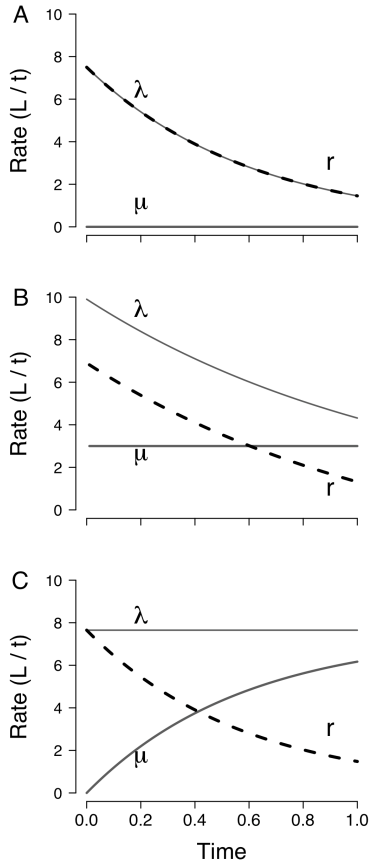

Figure 4.2 Models used to simulate phylogenies undergoing five-fold decline in net diversification rate (*r*; dashed line) attributable to (A) decreasing speciation rate (λ) and with an extinction rate (μ) equal to zero; (B) decreasing speciation, with extinction constant but greater than zero; and (C) speciation constant and extinction increasing through time. Rate through time curves are based on parameters given in Table 1 and were selected to result in a mean of 80 surviving lineages after 1.0 time units.

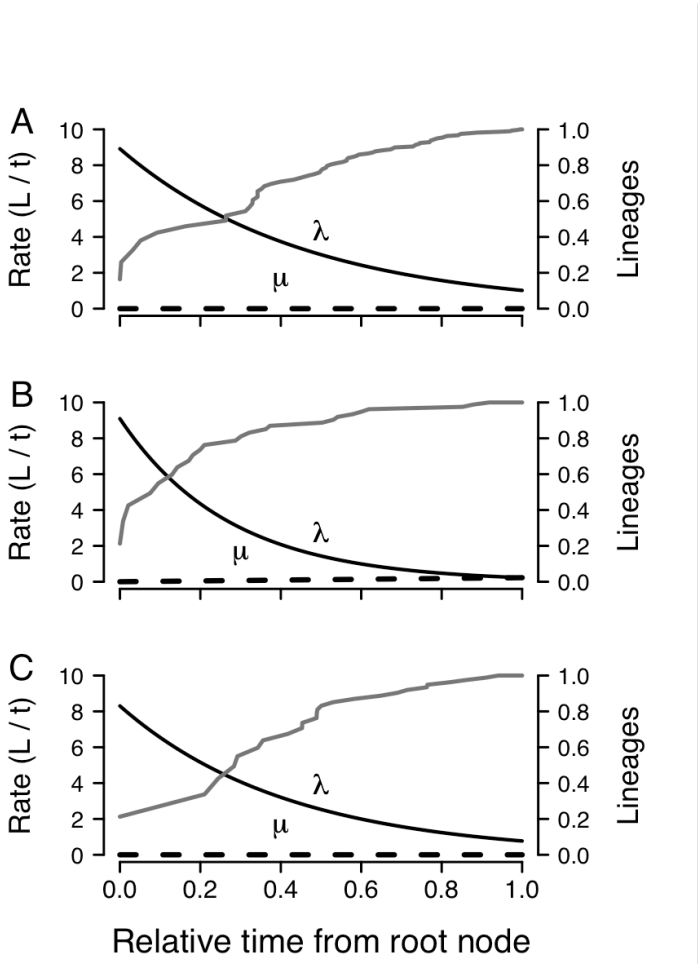

Figure 4.3. Maximum likelihood estimates of speciation rates (λ, solid line, decreasing) and extinction rates (μ, dashed line) under the BOTHVAR model for three phylogenies discussed in text: (A) Australian agamid lizards, (B) North American wood-warblers, and (C) Australo-Papuan pythons. The corresponding log-lineage through time curves (solid line, increasing) are included in each plot. Phylogenies were taken from original sources and rescaled to a basal divergence of 1.0 time units before the present. Rates are given in units of lineages per time unit. In each phylogeny, the extinction rate is inferred to have undergone minimal or no increase through time; in contrast, speciation rates consistently show a large decline. Speciation rates declined most rapidly in wood-warblers (B), as assessed by the slope of the speciation rate curve.

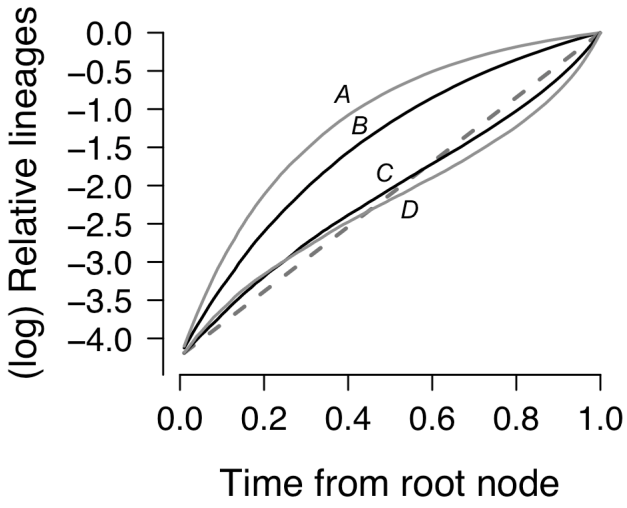

Figure 4.4. Mean log-lineage through time curves for phylogenies simulated under alternative diversification scenarios. (A) 15-fold decrease in net diversification rate (*r*) mediated by declining speciation rates, with no extinction; (B) 5-fold decrease in *r* mediated by declining speciation rates, with no extinction; (C) 5-fold decrease in *r* mediated by increasing extinction through time, with constant speciation; (D) 15-fold decrease in *r* mediated by increasing extinction, with constant speciation. Curves are based on 1000 simulations using diversification parameters given in Table 1 and were rescaled to a maximum of 1.0 lineages. Despite a large real decline in their net diversification rate, phylogenies generated under increasing extinction (C and D) show an characteristic upturn in the number of lineages towards the present, a pattern that is typically interpreted as stemming from increasing diversification through time or from high but constant relative extinction rates.

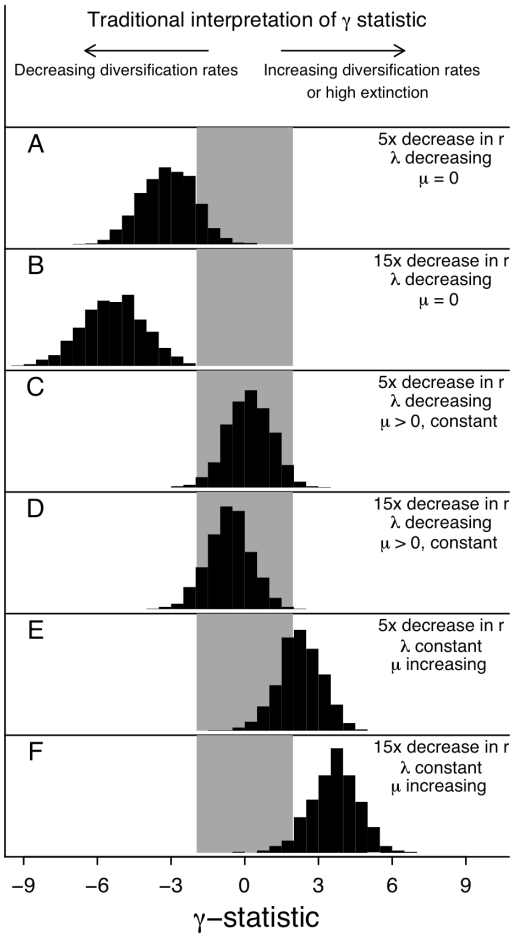

Figure 4.5 Distribution of γ statistic for phylogenies simulated with identical net diversification rates (*r*) but with different speciation (λ) and extinction (μ) parameterization. Gray region represents 95% confidence interval on the null hypothesis of constant diversification through time under the pure-birth ($\mu = 0$) model. (A) 5-fold and (B) 15-fold declines in *r* mediated by a decline in the speciation rate, with zero extinction. (C) 5-fold and (D) 15-fold declines in *r* mediated by declining speciation, but with high and constant rates of background extinction. (E) 5-fold and (F) 15-fold declines in *r* mediated by increasing extinction and constant speciation rates.

Table 4.1. Parameters used for simulating phylogenies undergoing temporal declines in the net diversification rate. Simulations were conducted for 1.0 time units, with parameter updates every 0.02 time units. Parameters were chosen to yield an average of 80 surviving lineages at the end of the simulation.

Model	*Rate decrease*	λ_0	k	μ_0	z
SPVAR	5x	7.5	1.64	0	-
SPVAR	15x	10.9	2.77	0	-
SPVAR	5x	9.9	0.83	3	-
SPVAR	15x	11.6	1.2	3	-
EXVAR	5x	7.65	-	7.5	1.73
EXVAR	15x	11.475	-	11.25	3.12

Table 2. – Results of fitting constant-rate (birth-death) and variable-rate (SPVAR, EXVAR, BOTHVAR) models to phylogenies of agamid lizards, wood-warblers, and pythons. Maximum log-likelihoods and ΔAIC scores (parentheses) are shown for each model, where the lowest ΔAIC indicates the best-fit model. For each phylogeny, the SPVAR model provided the best fit to the data. Δλ and Δμ indicate net change in speciation and extinction rates between time of the basal divergence and present day under SPVAR and EXVAR models, respectively.

Data	*Birth-death*	*SPVAR*	*EXVAR*	*BOTHVAR*	*Δλ*	*Δμ*
Agamids	207.9	217.7	207.9	217.7	-7.89	0
	(17.6)	(0)	(19.6)	(2.0)		
Warblers	42.1	52.9	42.1	52.9	-9.3	0
	(19.6)	(0)	(21.6)	(2.0)		
Pythons	49.2	53.3	49.2	53.3	-7.53	0
	(6.3)	(0)	(8.3)	(2.0)		

Why do models specifying temporal increases in extinction rates fail to fit these lineage accumulation curves better than a simple constant-rate birth-death model? Our analyses of simulated datasets with temporally declining net diversification rates provide a ready explanation for this phenomenon. Although both declining speciation and increasing extinction can yield identical net diversification rates through time, patterns of lineage accumulation vary dramatically between these competing models of diversification. When speciation rates decline through time with low background extinction, lineage-through-time (LTT) plots reveal a rapid rise in the number of lineages early in the history of the radiation (Figure 4.4A, B). This is the LTT relationship expected to result from declining diversification through time (Nee et al. 1992; Wollenberg et al. 1996; Pybus and Harvey 2000), and it is the pattern seen in the phylogenies for the three real taxonomic groups we analyzed as representatives of this common phenomenon. As expected, our simulations also show that the γ statistic becomes increasingly negative as the speciation rate decline becomes more severe (Figure 4.5A, B).

However, when changes in the net diversification rate are mediated solely by increasing extinction rates through time, reconstructed LTT curves bear little trace of the high diversification rates that were present in the early stages of a radiation (Figure 4.4C, D). Distributions of γ for such phylogenies indicate that the relative waiting times between successive speciation events (as inferred from the topology that includes only extant taxa) retain no signature of declining net diversification rates through time (Figure 4.5E, F). When phylogenies are simulated under the EXVAR model, we found that the largest declines in diversification rates yield the largest values of γ (Figure 4.5F); such positive values of γ are traditionally interpreted as consistent with increasing diversification through time or high extinction (e.g., Barraclough and Vogler 2002; Linder et al. 2003). These high γ values under high

extinction are almost certainly due to the 'pull of the present' (Nee et al. 1994a, b), whereby high relative extinction rates – the ratio of extinction to speciation - create an apparent excess of recently diverged lineages in reconstructed phylogenies. This phenomenon has been discussed previously as a potentially confounding issue in diversification analyses (Nee 2001; Rabosky 2006b), because high relative extinction rates can create the spurious impression of increasing diversification through time, even when rates have not changed. Here we find that this effect is strong enough to overcome even massive declines in net diversification rate, potentially leading researchers to infer a temporal increase in net diversification rates in situations where net diversification is actually declining via increasing extinction.

These results imply that traditional interpretations of LTT plots and associated test statistics may be naive if the potential role of extinction is neglected. An apparent excess of recently diverged lineages in LTT plots is typically interpreted as consistent with increasing diversification through time or high relative extinction rates (Barraclough and Vogler 2002; Turgeon et al. 2005; Rabosky 2006b; Roelants et al. 2006). Our results indicate that that declining net diversification through time could yield similar patterns of lineage accumulation in reconstructed phylogenies, if the decline is driven by increasing extinction rates.

We also found that high but constant extinction rates erode the signature of explosive-early speciation. We expected to observe some reduction in our ability to detect temporally declining speciation under high background extinction, as the 'pull of the present' would partially offset the rapid rise in lineages at the start of the radiation. However, we were unprepared for the observation that the distribution of γ from even a 15-fold decline in the net diversification rate was virtually indistinguishable from a constant rate diversification process (Figure 4.5C, D) under high but constant extinction rates. This result suggests that, when background

extinction rates have been high, even large declines in the rate of speciation through time will be difficult to detect using phylogenies of extant taxa only.

The power of extinction, whether constant or variable, to influence LTT plot-based inferences about diversification rates raises an important question: how often will real-world extinction rates be high relative to speciation rates? Evidence from the fossil record supports the view that relative extinction rates are generally high (e.g., Stanley 1979; Stanley et al. 1988; Gilinsky 1994; Newman and Sibani 1999), and we are unaware of any evidence that extant clades have diversified in the absence of extinction. In the case of mammals, for example, Alroy (1996) found mean per-genus and per-lineage relative extinction rates of 0.90 and 0.91, respectively, across 55 1.0 million year intervals during the Cenozoic. In Gilinsky's (1994) tabulation of familial origination and extinction rates in marine invertebrates, nearly two-thirds of all orders have had relative extinction rates in excess of 0.8 (91/137). This raises a conundrum: the fossil record suggests that real clades will often evolve under conditions that should make it difficult to ever detect temporally declining speciation rates from LTT analyses of extant species, yet a large number of empirical studies have documented exactly that pattern across a diverse range of taxonomic groups. How can it be that so many radiations provide strong LTT evidence for explosive-early diversification?

We suggest two possible solutions to this seemingly paradoxical observation. The first is simply that fossil-derived relative extinction rates do not apply to phylogenies of extant taxa. This could be the case if relative extinction rates are highly conserved among closely-related taxa (e.g., Heard and Mooers 2000). For example, mammals are characterized by high relative extinction rates (Alroy 1996), but it is not clear whether these overall rates apply to the subset of mammalian lineages that have actually survived to the present and are hence available for LTT-based comparisons. A number of mammalian subclades have gone extinct entirely (Bininda-Emonds et al.

2007), and relative extinction rates for at least some clades that survived to the present are somewhat lower than Alroy's overall (1996) estimate (Muñoz-Duran 2003).

A second possibility is that explosive-early diversification is often an artifact of methodological biases stemming from taxon sampling issues or the methodologies used to generate phylogenetic trees for LTT comparisons. Current LTT approaches assume that extant taxa of equivalent biological rank (e.g., "species") are comprehensively sampled, and that the temporal distribution of nodes (i.e., branch lengths) in their phylogeny is not biased with respect to age. Incomplete taxon sampling will result in a spurious decline in speciation rates as inferred from reconstructed phylogenies (Nee et al. 1994a; Pybus and Harvey 2000). We recognize that many clades will contain unrecognized or unsampled lineages whose exclusion could bias diversification analyses; this may be a particularly prevalent issue in clades where incipiently divergent phylogeographic lineages are not recognized and included as incipient species. Failure to use an appropriate model of sequence evolution may also lead to disproportionate compression of early branches in phylogenetic trees (Revell et al. 2005), thus creating the impression of decreasing speciation through time. Similar problems may also be associated with different algorithms for constructing ultrametric trees (e.g., Ruber and Zardoya 2005). However, these are methodological artifacts of the phylogeny reconstruction process, not of the LTT approaches applied to those phylogenies.

The analytical framework described here for modeling speciation and extinction rates that vary continuously through time should be applicable to a range of problems involving temporal and among-lineage variation in diversification rates. The basic model for time-varying diversification rates (eqns 2 - 6) can be modified to allow speciation and extinction probabilities to vary among lineages, perhaps as a function of species trait values (e.g., Paradis 2005). The advantages of using a general

framework based on the birth-death process are twofold. First, in contrast to simple parametric and non-parametric diversification test statistics (Wollenberg et al. 1996; Paradis 1998; Pybus and Harvey 2000), our model-fitting approach provides biologically meaningful parameter estimates and – as demonstrated here – can be used to infer changes in both speciation and extinction rates through time. This distinguishes our approach from survival analysis (Paradis 1997), which can accommodate continuous-time variation in speciation rates but is limited by its explicit assumption of zero extinction (Felsenstein 2004).

A second advantage of the present approach is that it provides researchers the power to address specific macroevolutionary questions with a large range of biologically relevant diversification models. There is nothing special about the models for time-varying diversification we selected for this study; we chose them for their simplicity (3 or 4 parameters), flexibility (we could model non-linear changes in speciation and extinction rates), and because they permitted us to address our focal question. The computational tools for numerical integration and optimization available through R, MATLAB, and other analysis platforms make it possible to fit diversification models that are more complex than those discussed in the present paper. Of course, it is always the case that models can only approximate evolutionary processes, and matching a model to a particular question is not a trivial undertaking (Bolker 2008). Knowledge of dubious quality is gained when one poorly formulated model is found to fit the data better than another poorly formulated model.

In summary, we found that explosive-early radiations, as inferred from molecular phylogenies of extant taxa, can only be explained by temporal declines in speciation rates and not by increasing extinction rates through time. Some theoretical work suggests that extinction rates should increase through time during evolutionary radiations as a function of mean population sizes or per-capita resource availability

(e.g., Levinton 1979; Hubbell 2000). If this occurs, it is unlikely to leave a signature of early, rapid diversification in molecular phylogenies. To the extent that patterns of lineage accumulation observed in empirical datasets are not artifacts of biased branch-length reconstruction or incomplete taxon sampling, our results suggest that many clades appear to undergo rapid diversification early in their history because speciation but not extinction rates have changed over the histories of those groups.

CHAPTER 5

HERITABILITY OF EXTINCTION RATES LINKS DIVERSIFICATION PATTERNS IN MOLECULAR PHYLOGENIES AND FOSSILS

Abstract

Time-calibrated molecular phylogenies provide a valuable window into the tempo and mode of species diversification, especially for the large number of groups that lack adequate fossil records. Molecular phylogenetic data frequently suggest an initial 'explosive speciation' phase, leading to widespread speculation that ecological niche-filling processes might govern the dynamics of species diversification during evolutionary radiations. However, these patterns are difficult to reconcile with the fossil record: the fossil record strongly suggests that extinction rates have been high relative to speciation rates, but such elevated background extinction should erase the signal of early, rapid speciation from molecular phylogenies. For this reason, extinction rates in molecular phylogenies are frequently estimated as zero under the widely-used birth-death model. Here I construct a simple model that combines phylogenetically-patterned extinction with pulsed turnover dynamics and constant diversity through time. Using approximate Bayesian methods, I show that heritable extinction can easily explain the phenomenon of explosive early diversification, even when net diversification rates are equal to zero. Several assumptions of the model are more consistent with both the fossil record and neontological data than the standard birth-death model and it may thus represent a viable alternative interpretation of phylogenetic diversification patterns. These results suggest that variation in the absolute rate of lineage turnover through time, in conjunction with phylogenetically non-random extinction, may underlie the apparent density-dependent speciation observed in molecular phylogenies.

Introduction

Molecular phylogenetic studies can potentially complement fossil-based analyses of evolutionary radiations, because time-calibrated molecular phylogenies provide information about the timing of speciation events in groups for which minimal fossil data are available. The ease of obtaining DNA sequence data has led to a rapid increase in the availability of time-calibrated phylogenetic trees, and many studies have used these data to infer patterns of species diversification through time (Barraclough and Vogler, 2002; Harmon et al., 2003; Nee et al., 1992). Perhaps the most surprising finding from these studies is that many, if not most, suggest an initial burst of lineage accumulation early in the history of evolutionary radiations (McPeek, 2008; Phillimore and Price, 2008; Ruber and Zardoya, 2005). This pattern has been frequently been interpreted as density-dependent speciation mediated by ecological opportunity, whereby rapid diversification is facilitated by an abundance of resources and paucity of competing species (Nee et al., 1992; Phillimore and Price, 2008; Price, 2008; Rabosky and Lovette, 2008a). Under this model, speciation rates are high initially, but subsequently decline in conjunction with the rise of species diversity within a particular ecological or biogeographic theatre.

However, extinction shapes phylogenetic trees by 'pulling' nodes close to the present (Fig. 5.1) and thus obscures the signal of early, rapid diversification (Rabosky and Lovette, 2008b; Rabosky and Lovette, 2008c). This occurs because high background extinction changes the age structure of lineages that survive to the present to be observed. If the speciation rate λ is high relative to the extinction rate μ, then many lineages will have been around for a comparatively long period of time. However, if μ approaches λ, the turnover rate of lineages will be much higher, and most lineages will have been in existence for a relatively brief period of time. A corollary of this is that early rapid diversification is more difficult to detect in

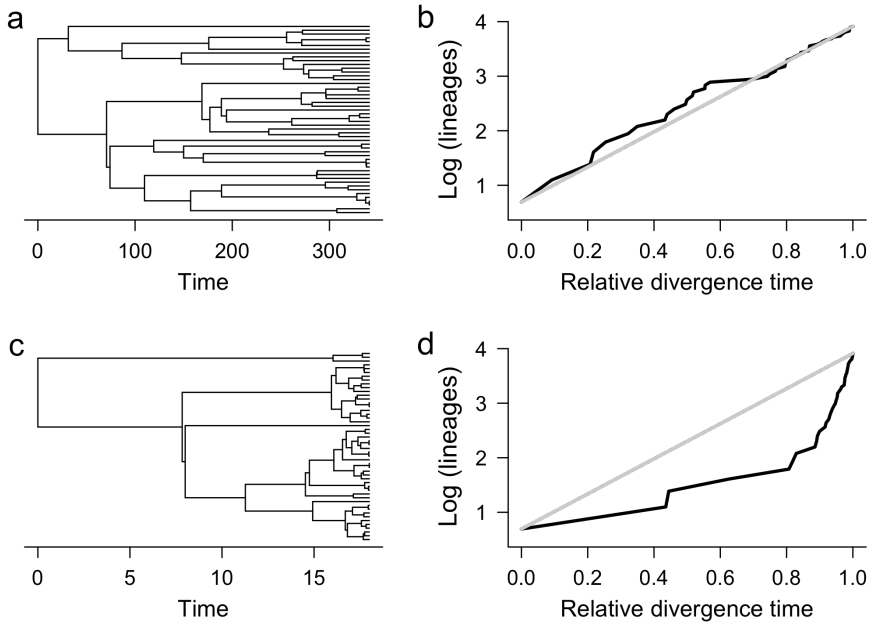

Figure 5.1. High extinction rates create the impression of accelerating diversification through time and eliminate the signal of early branching lineages. (a) Representative phylogenetic tree simulated under pure-birth process, with constant speciation (λ) through time with extinction (μ) equal to zero. (b) Lineage accumulation curve for tree shown in (a); gray line indicates expected lineage accumulation curve under pure birth process. (c) Phylogenetic tree simulated under constant rate birth-death process, with high relative extinction rate ($\mu / \lambda = 0.99$), and (d) corresponding lineage accumulation curve. Net diversification rate ($\lambda - \mu$) was identical for both trees (.01 lineages/time unit), and both trees contain $N = 50$ surviving lineages, but the pure birth tree is much deeper (~350 time units) than the high extinction tree (~18 time units). Note also the 'pull of the present' in the high extinction lineage accumulation curve (d), which shows an apparent rapid increase in the rate of lineage accumulation towards the present, even through the tree was simulated under constant rate model of diversification.

molecular phylogenies if extinction rates have been high (Rabosky and Lovette, 2008b). This underlies the observation that extinction rates estimated from molecular phylogenies are typically close to zero (Nee, 2006; Purvis, 2008; Rabosky and Lovette, 2008b; Weir, 2006). Indeed, under the widely-used birth-death model (Nee et al. 1994), phylogenies showing an excess of early speciation events almost invariably appear to have trivial extinction (Rabosky and Lovette, 2008b).

These results directly conflict with the fossil record (Bokma, 2008), which overwhelmingly suggests that extinction rates have been high relative to speciation rates, in every group for which appropriate data are available (Alroy, 1996; Alroy, 2008; Alroy, 2009; Gilinsky, 1994; Stanley, 1979). Studies of paleodiversity through time lend further support to this idea, because a consistent and substantial excess of speciation relative to extinction would lead to exponentially increasing diversity for many groups. Yet this is clearly not the case: studies that have investigated sampling-standardized diversity through time have found no evidence for recent exponential increases in diversity (Alroy et al., 2008; Rabosky and Sorhannus, 2009) and many groups appear to be characterized by approximate constancy of diversity through time (Alroy, 2000; Jaramillo et al., 2006). Moreover, clade age is unrelated to species richness in many higher taxa, a result that implies ecological limits on clade diversity through time (Rabosky, 2009; Ricklefs, 2006; Ricklefs et al., 2007).

A more subtle problem with putative 'density dependent' interpretations of these patterns is that, for most systems, there is little historical evidence for conditions of high ecological opportunity and empty ecological space that could have triggered rapid speciation. It is certainly the case that diversification on isolated islands may be triggered by ecological opportunity (Schluter, 2000) and that conditions favoring explosive ecological and species diversification might follow the profound biotic upheaval of mass extinctions (Foote, 1996; Sepkoski, 1998). But it is much more

difficult to see how so many radiations of taxa in ecologically complex continental systems might have experienced much greater niche availability early in their history. There is little evidence for massive clearing of ecological space by extinction over timescales consistent with molecular phylogenetic analyses of diversification (Alroy, 2009; Ricklefs, 2007). Even if Pleistocene climate oscillations resulted in recent extinction pulses in the temperate zone, many groups from tropical regions that would have been minimally affected by these events have often undergone explosive early diversification (Weir, 2006).

While biases associated with taxon sampling (Purvis et al., 2008; Pybus and Harvey, 2000) and tree construction (Revell et al., 2005) might partially account for apparent conflict between molecular and fossil data, it is also possible that the conflict is an artifact of the standard birth-death model used to draw inferences about species diversification rates from molecular data. Here, I construct a simple alternative to the birth-death model that assumes (i) constant diversity through time, (ii) heritable extinction tendencies, and (iii) a balanced speciation-extinction process with zero net diversification. The effects of heritable extinction rates on lineage accumulation curves has not been studied, although a number of studies have documented both heritability of diversification rates (Davies et al., 2004; Savolainen et al., 2002) as well as phylogenetic clustering of extinction and extinction risk (Purvis et al., 2005; Vamosi and Wilson, 2008). In the new model, each extinction event results in an immediate replacement, as in Hubbell's neutral community model (Hubbell, 2001).

I apply the model to a recent phylogeny of *Dendroica* wood-warblers, a well-studied group of North American songbirds that appear to have undergone a density-dependent decline in the rate of speciation with trivial background extinction (Rabosky and Lovette, 2008a; Rabosky and Lovette, 2008b). Because no likelihood function can presently be specified for the model, I used approximate Bayesian

computation to infer posterior distributions of parameters and to evaluate absolute model fit.

I show that a simple turnover pulse with phylogenetically patterned extinction can account for speciation-extinction dynamics observed in molecular phylogenies, with ecological implications that contrast sharply with inferences based on the birth-death model.

Methods

Heritable Extinction with Pulsed Turnover (HEPT) model

The birth-death model as typically used for inference on species diversification rates assumes per-lineage rates of speciation (λ) and extinction (μ) in a clade that grows from an initial diversity of $n = 2$ lineages (Nee et al., 1994). In contrast, the heritable extinction with pulsed turnover (HEPT) model I consider is a Moran process with fixed diversity through time and parameters $\theta = (x_c , \kappa, r_1, r_2, \omega)$. Speciation and extinction rates are equal and events occur simultaneously, such that the extinction of one species immediately results in speciation by another.

At some time ($x_c + \kappa$) before present, per-lineage speciation and extinction rates r_1 shift to a new rate r_2 (the 'turnover pulse' phase), which is maintained for 2κ time units, until time $x_c - \kappa$, whereupon r_2 shifts back to r_1 (Fig. 5.2). A crucial feature of this model is the heritability of extinction, which is specified by ω. If a particular species goes extinct, the next extinction event will eliminate a randomly drawn member of its sister taxon with probability ω. Otherwise, a taxon goes extinct at random from the full set of N taxa with probability $1 - \omega$. Note that $\omega = 0$ corresponds to the simple Moran process, at least if $r_1 = r_2$.

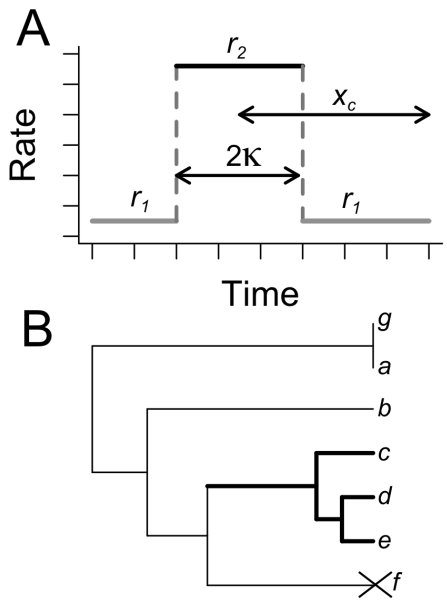

Figure 5.2. Parameters in heritable extinction with pulsed turnover (HEPT) model. (a) speciation and extinction rates r_1 shift to a new rate r_2 at some time $x_c + \kappa$ time units before the present. Lineage turnover with rate r_2 continues until time $x_c - \kappa$, at which point rates shift back to r_1. Note that speciation and extinction rates are equal. (b) The ω parameter controls the extent to which extinction events are phylogenetically structured. Here, lineage *f* has become extinct and lineage *g* has just originated. The next extinction event will eliminate a species from *f*'s sister clade (bold lines; taxa *c*, *d*, *e*) with probability ω, or it will eliminate a species at random from the tree (*a*, *b*, *c*, *d*, *e*, *g*) with probability $(1 - \omega)$. The total probability that lineage *d* would be next to go extinct is thus $\omega/3 + (1 - \omega)/6$.

Fitting the model to data with approximate Bayesian computation

I used approximate Markov Chain Monte Carlo (MCMC) (Marjoram et al., 2003; Plagnol and Tavare, 2004) to fit the HEPT model to a complete, species-level phylogeny for the *Dendroica* subgroup of North American wood-warblers (Rabosky and Lovette, 2008a). The tree contains all 25 species in the *Dendroica* group that occur within North America and was constructed under a Bayesian relaxed-clock approach from approximately 9500 bp of combined nuclear and mitochondrial DNA. Rabosky and Lovette (2008a) reported that *Dendroica* has undergone a severe decline in the rate of diversification through time, consistent with several previous studies that have also reported declining diversification through time in this group (Lovette and Bermingham, 1999; Phillimore and Price, 2008; Rabosky and Lovette, 2008b). I scaled the basal divergence of the *Dendroica* tree to 5 million years (my) before present (Lovette and Bermingham, 1999).

To obtain independent samples from the joint posterior distribution of HEPT model parameters, I used an approximate MCMC approach (Marjoram et al., 2003; Plagnol and Tavare, 2004) that does not require calculation of likelihoods (Marjoram et al., 2003). The approach requires only that we have one or more sufficient summary statistics that can be used to compare the observed data to data simulated under the model given a set of parameters. The approach is essentially a variant of rejection sampling: parameters are sampled from proposal distributions, and data are simulated under the HEPT model with the sampled parameters. We then compute summary statistics that describe the match between the observed data and the simulated data. Parameter sets that can reproduce phylogenetic patterns consistent with the observed data are retained, while those that cannot are rejected.

In this case, the depth (or crown age) of simulated phylogenetic trees (L) and the γ-statistic (Pybus and Harvey, 2000) provide convenient summary statistics that

describe the match between the observed wood-warbler tree and simulated trees. The γ statistic is a measure of the extent to which speciation events in a molecular phylogeny are clustered at the base or tips of a tree; $\gamma < 0$ implies clustering near the base of the tree and decelerating speciation through time (Pybus and Harvey, 2000; Rabosky and Lovette, 2008b).

The rejection criterion in an approximate Bayesian framework depends on the match between the observed data (D) and the simulated data (D'), as determined by the difference in summary statistics. We thus generate a candidate set of parameters θ_c and simulate data D' under the model. The parameters θ_c are retained if the distance $\rho(D', D)$ is less than some threshold ε. Parameter proposals could only be accepted if simulated trees approximately matched the observed tree in two summary statistics: the crown age, or tree depth (5 my), and the value of the γ statistic (-3.48). To generate phylogenies under the HEPT model, a candidate tree was first generated under a pure birth model with $\lambda = 0.1$. This tree was then evolved under the HEPT model for 20 my. The rejection criteria (discussed below) ensured that simulated trees retaining any signal of this initial pure birth tree were not accepted; increasing the simulation duration to 100 my did not change the results, but dramatically increased computational time.

Initial parameters for each independent chain were sampled randomly from prior distributions $\pi(\cdot)$. Priors were uniform for all parameters [r_1 and r_2 (0, 30); κ (0, 5), and ω (0, 1)] except x_c, which received a normal (mean = 5, sd = 2) prior. The MCMC sampler then iterated through the following steps, until a target number of generations had been reached (Marjoram et al., 2003):

S1: If at state θ, propose a move to state θ' according to a transition kernel $q(\theta \rightarrow \theta')$

S2: Generate data D' under the HEPT model with parameters θ'

S3: Compute $\rho(D', D)$. If $\rho(D', D) \leq \varepsilon$, go to S4. Otherwise, reject the proposal and return to S1.

S4: Compute an approximation of the Hastings ratio as

$$h = \min\left(1, \frac{\pi(\theta')q(\theta' \rightarrow \theta)}{\pi(\theta)q(\theta \rightarrow \theta')}\right) \quad (5.1)$$

where $\pi(\theta') / \pi(\theta)$ is the prior ratio and $q(\theta' \rightarrow \theta) / q(\theta \rightarrow \theta')$ is the ratio of transition kernels between states.

S5: Accept the proposal θ' with probability h; otherwise, stay at θ. Return to S1.

Here, the ratio of transition kernels is always equal to one. The criterion $\rho(D', D)$ is simply the absolute difference in γ and L between the simulated and observed data, scaled by the absolute value of the observed γ and L. Thus, for gamma, this was simply $| \gamma_{SIM} - \gamma_{OBS} | / | \gamma_{OBS} |$. After experimenting with a range of ε between 0.05 and 0.5, I found that $\varepsilon = 0.2$ performed well for both γ and L. A proposed state could thus only be accepted if the scaled differences in both γ and L were less than ε. I ran 10 chains for 100,000 generations each, sampling parameters every 100 generations.

I repeated these analyses assuming fixed ω values. Multiple chains were run for scenarios corresponding to low ($\omega = 0.1$), moderate ($\omega = 0.5$), and high ($\omega = 1.0$) extinction heritability. Together with the simulations in which ω was treated as a free parameter, this gave a total of four evolutionary scenarios analyzed with approximate MCMC. All simulation and analyses were conducted in the R programming environment, with some code borrowed from the Ape (Paradis et al., 2004), Geiger (Harmon et al., 2008), and Laser (Rabosky, 2006) packages.

Convergence analysis and diagnostics

Each MCMC chain was initiated with a randomly drawn set of parameters that satisfied $\rho(D', D) \leq \varepsilon$. There was thus no need to include a burn-in phase as is typical with standard MCMC, because the chain was already sampling from the target distribution. To assess convergence, I calculated Gelman and Rubin's scale reduction factor (Gelman and Rubin, 1992). This statistic essentially compares within- and between-chain variances for each parameter and estimates the factor by which the scale parameter of the estimated posterior densities for each parameter would shrink were the chains to run to infinity. There are no absolute guidelines for interpreting this statistic, but values less than 1.05 are typically taken to reflect convergence (Gelman et al., 2003). Prior to computing the Gelman-Rubin statistic, data were transformed to improve normality.

Lineage accumulation curves

I assessed whether lineage accumulation curves generated under the HEPT model showed evidence for high or low extinction when analyzed with several variants of the birth-death model. In this case, the simulation model specifies $\lambda = \mu$, so a failure to recover high extinction relative to speciation implies a failure of the birth-death model to reconstruct a basic feature of the underlying evolutionary process. For each value of ω, mean log-lineage accumulation curves were tabulated from the sample of simulated phylogenies obtained during MCMC. These phylogenies necessarily satisfied $\rho(D', D) \leq 0.2$ for both γ and L. The first model fitted to the mean lineage accumulation curves was simply a constant-rate birth-death process (Nee et al., 1994). The second was the BOTHVAR model from Rabosky and Lovette (Rabosky and Lovette, 2008b) which allows simultaneous, independent changes in both speciation and extinction rates, such that the net diversification is modeled as

(5.2) $r(t) = \lambda(t) - \mu(t)$,

where

(5.3) $\lambda(t) = \lambda_0 \exp(-kt)$

and

(5.4) $\mu(t) = \mu_0 (1 - \exp(-kt))$.

I estimated the time-integrated extinction fraction μ / λ as

(5.5) $$\frac{\mu}{\lambda} = \frac{\int \mu(t) dt}{\int \lambda(t) dt}$$

integrated from the time of the basal divergence to the present. The final model specified exponentially declining speciation and extinction rates, where the extinction rate is a constant fraction υ of the speciation rate, leading to

(5.6) $\lambda(t) = \lambda_0 \exp(-kt)$

and

(5.7) $\mu(t) = \upsilon \lambda(t)$.

Posterior predictive simulations

The preceding analyses tell us little about how well the HEPT model fits the wood-warbler data. To determine whether the model is capable of recovering the major features of wood-warbler diversification, I assessed absolute model fit by simulating phylogenies under parameters sampled from the joint posterior distribution of θ conditional on several values of ω. For each simulated dataset, I then calculated the summary statistics γ and L, which indicate (i) whether the simulated dataset shows evidence for an apparent slowdown in the rate of diversification through time, and (ii) whether the simulated tree is of the same age as the wood-warbler radiation. Here, I was simply using parametric simulation to determine the adequacy of the HEPT model, given the joint posterior distribution of parameters $f(\theta \mid D, \omega, M)$. Formally, we are interested in the distribution of one or more summary statistics Z' given the summary statistic Z observed for the real data,

$$\Pr(Z' \mid Z, M, \omega) = \int \Pr(Z' \mid \theta, M, \omega) \Pr(\theta \mid D, M, \omega) d\theta \tag{5.8}$$

In practice, parameters were sampled at random and with replacement from the posterior distribution of parameters $f(\theta \mid D, \omega, M)$. Each random draw consisted of selecting x_c, κ, r_1, and r_2 from a single generation of the pooled MCMC chains, thus ensuring that parameters were sampled in proportion to their joint posterior probability. A single simulation of the HEPT model was performed for each sampled set of parameters, and the distribution of Z' was tabulated from 50,000 such random draws per ω model. I evaluated absolute model fit under scenarios of high ($\omega = 1.0$), moderate ($\omega = 0.5$), and low ($\omega = 0.1$) extinction heritability. I found the posterior distributions $f(\theta \mid D, \omega, M)$ of x_c, κ, r_1, and r_2 conditional on these values of ω by running five MCMC chains per ω.

Results

HEPT model

Analysis of 10 independent MCMC chains (Table 5.1) with the Gelman-Rubin statistic (Gelman et al., 2003) suggested that all chains converged on the same target distribution (Table 5.2). Figure 5.3 shows posterior distributions for ω, κ, and rate parameters inferred using approximate MCMC. The posterior distribution of ω suggests that patterns of wood-warbler diversification through time are most consistent with phylogenetically structured extinction under the HEPT model (Fig. 5.3a). The distribution of κ and the ratio of turnover rates (r_2 / r_1) implies that a broad range of turnover conditions, both in terms of rates and duration, may also be consistent with the observed data (Fig. 5.3b, c). However, there is a strong negative relationship between these parameters: turnover pulses of short duration require much greater turnover rates, and vice-versa (Fig. 5.3d).

Per-lineage estimates of speciation and extinction were summarized from the results for each ω category (ω variable; $\omega = 1.0$; $\omega = 0.5$; $\omega = 0.1$); median values of each posterior distribution, as well as 2.5% and 97.5% quantiles, are shown in Table 5.3. The magnitude of the turnover pulse (r_2) is negatively correlated with ω. When extinction events are not phylogenetically clustered ($\omega = 0.1$), increased speciation and extinction rates are required to generate phylogenies consistent with the observed data.

Lineage accumulation curves

Visual comparison of mean log-lineage accumulation curves obtained for different ω values clearly indicate that high extinction heritability provides the best match to the observed pattern of lineage accumulation through time for the wood-

Table 5.1. Run and sampling statistics for approximate MCMC analyses.

ω model	*Runs*	*gens/run*	*Sample frequency*	*Acceptance rate*	*Scale reduction factor*[1]
ω variable	10	100000	100	0.053	1.01
ω = 0.1	5	100000	100	0.014	1.05
ω = 0.5	5	100000	100	0.034	1.01
ω = 1.0	5	100000	100	0.096	1

[1]Multivariate proportional scale reduction factor (Brooks and Gelman 1997).

Table 5.2. Gelman-Rubin (Gelman & Rubin 1992) proportional scale reduction factor estimates, with 0.975 percentiles, for model parameters (all parameters transformed to improve normality).

Parameter	*ω variable*	*ω = 0.1*	*ω = 0.5*	*ω = 1.0*
ω	1.00 (1.01)	NA	NA	NA
x_c	1.00 (1.01)	1.00 (1.01)	1.01 (1.02)	1.00 (1.01)
κ	1.01 (1.02)	1.02 (1.05)	1.01 (1.03)	1.00 (1.01)
r1	1.01 (1.01)	1.07 (1.16)	1.01 (1.03)	1.00 (1.00)
r2	1.01 (1.03)	1.02 (1.04)	1.01 (1.02)	1.00 (1.01)

Table 5.3. Speciation and extinction rates for wood-warblers under the HEPT model during periods of low (r_1) and high (r_2) turnover. Shown are medians from the posterior distribution, as well as 2.5% and 97.5% quantiles. Note that speciation and extinction rates are equal under the HEPT model. Rates are in lineages/my assuming the basal wood-warbler divergence occurred 5 my before present.

Simulation model	*r_1 (base rate)*	*r_2 ('turnover pulse')*
ω as free parameter	0.18 (0.04, 0.36)	3.57 (1.13, 22.57)
ω = 0.1	0.18 (0.05, 0.35)	2.97 (0.82, 21.30)
ω = 0.5	0.17 (0.03, 0.34)	3.99 (1.40, 24.24)
ω = 0.1	0.17 (0.02, 0.33)	4.44 (1.66, 24.53)

warbler data (Fig. 5.4). However, even low extinction heritability yielded a 'concave down' lineage accumulation curve consistent with declining diversification through time. For all three variants of the birth-death model, extinction was inferred to be a trivial fraction of the speciation rate (Table 5.4), in spite of the fact that the simulation model specified exact equivalence between these parameters.

Posterior predictive simulations

When extinction shows high phylogenetic structure ($\omega = 1.0$), posterior predictive distributions of both γ and L are centered on the observed values for the wood-warbler tree (Fig. 5.5a, b), indicating that this model fits the data well. The model does not perform as well with intermediate extinction heritability ($\omega = 0.5$) and appears to perform poorly with low extinction heritability. For $\omega = 0.1$ and 0.5, both γ and L show bimodal distributions (Fig. 5.5a, b), and these values are correlated (Fig. 5.5c): trees that are young (3-8 Ma) show lineage accumulation patterns consistent with declining diversification rates ($\gamma < 0$). However, trees that are far older than the wood-warbler tree show $\gamma > 0$, which is inconsistent with temporally declining diversification.

This bimodality in tree depth and γ is due solely to the fact that some trees have a long waiting time between the initial bifurcation in the phylogenies and the next speciation event in the reconstructed tree. This was immediately obvious from visual inspection of both simulated trees and corresponding lineage-through-time plots (Fig. 5.6). To demonstrate this, I computed γ for each tree from the $\omega = 0.1$ posterior predictive simulations, but ignored the basal divergence in the tree (Fig. 5.6). When this basal bifurcation is ignored, the distribution of γ strongly suggests temporally declining diversification of similar magnitude to that observed in wood-warblers (Fig. 5.5d), even though these trees were simulated under $\omega = 0.1$. Thus, even limited

diversity models with phylogenetically unstructured extinction can generate lineage-through-time patterns consistent with temporally declining or density-dependent diversification. However, this pattern is more readily apparent with heritable extinction, which tends to eliminate the signal of diversification that occurred prior to the turnover pulse. This leaves the impression that a single clade has undergone rapid diversification followed by a temporal decline in rates.

Discussion

These results have broad implications for the interpretation of species diversification patterns as inferred from molecular phylogenies. I demonstrated that a simple limited diversity model, with equivalent speciation and extinction rates and zero net diversification through time, can generate patterns of lineage accumulation in molecular phylogenies that suggest temporally declining, density dependent diversification rates. This apparent slowdown in the rate of speciation is driven solely by variation in the rate of lineage turnover through time in conjunction with phylogenetically patterned extinction. Such a pulsed turnover model can account for patterns of lineage accumulation observed in North American wood warblers, a group believed to have undergone explosive speciation early in its history (Rabosky and Lovette, 2008a; Rabosky and Lovette, 2008b). Posterior predictive simulations under the HEPT model consistently recovered major features of wood-warbler diversification, including the rapid accumulation of lineages during the early stage of the radiation.

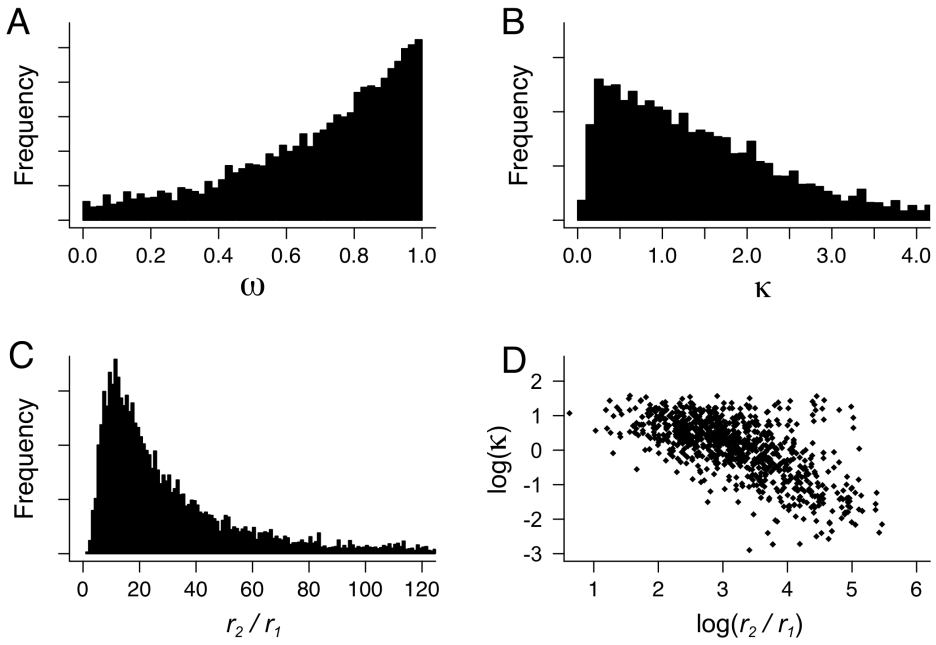

Figure 5.3. Posterior distributions of parameters in the HEPT model sampled using approximate MCMC including (a) ω, (b) κ, and (c) the ratio of turnover rates r_2 and r_1. (d) The ratio of turnover rates and the radius of the turnover pulse κ are negatively correlated; shorter turnover windows require greater turnover rates to yield appropriate tree depth and γ statistics. Posterior distribution of ω (a) suggests that, under the HEPT model, extinction events in wood-warblers are phylogenetically clustered.

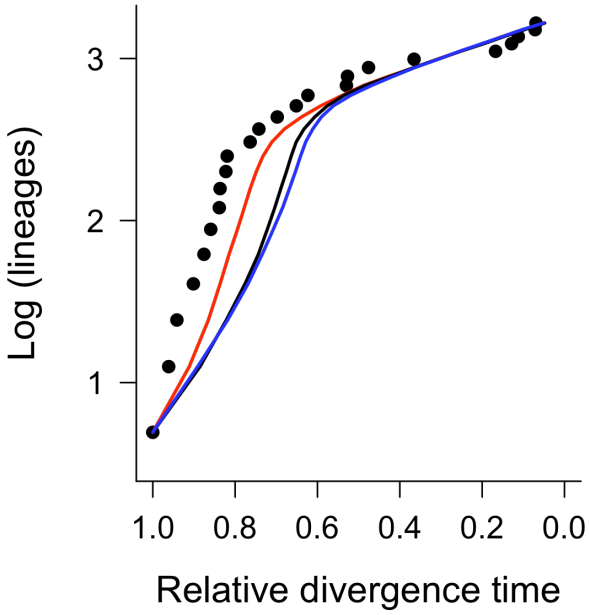

Figure 5.4. Log-lineage accumulation curve for *Dendroica* wood-warblers (Rabosky and Lovette, 2008a) (black circles) and mean log-lineage accumulation curves for trees sampled using approximate MCMC under $\omega = 1.0$ (red), $\omega = 0.5$ (black), and $\omega = 0.1$ (blue).

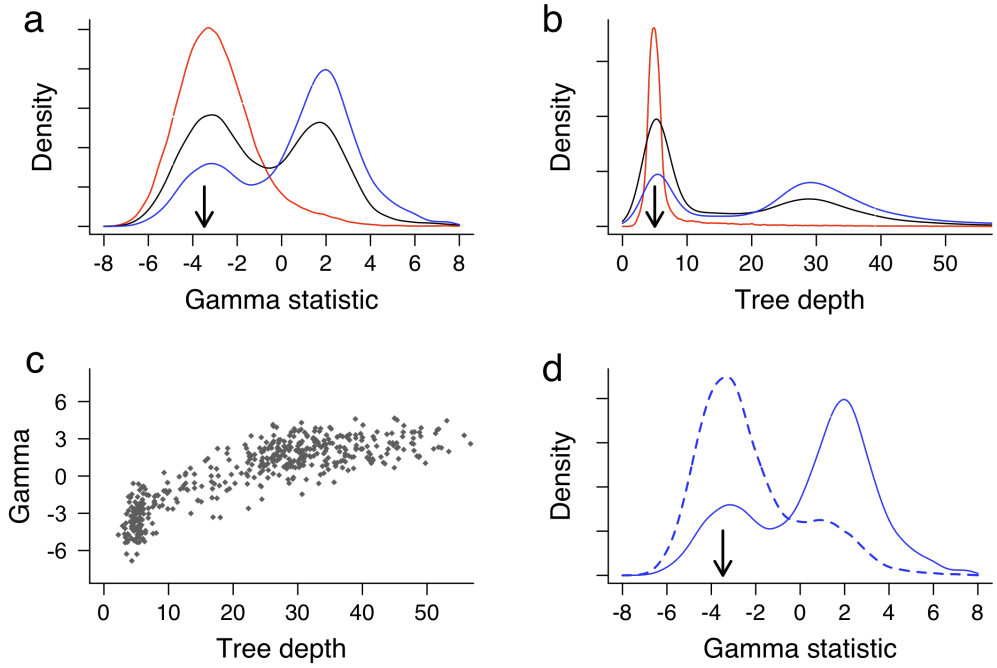

Figure 5.5. Posterior predictive simulations. Distribution of γ-statistic (a) and tree depth (b) from 50,000 simulations with parameters inferred under $\omega = 1.0$ (red), $\omega = 0.5$ (black), and $\omega = 0.1$ (blue). Arrows indicate observed values for *Dendroica* wood-warblers. (c) Relationship between γ and tree depth for parameters inferred with $\omega = 0.1$; trees that are older than the wood-warbler tree also tend to have positive γ values. (d) Distribution of γ for $\omega = 0.1$ while ignoring (dashed) and including (solid) the basal bifurcation in the tree. Ignoring the earliest speciation event gives a distribution of γ very similar to the wood-warbler data (Fig. 5.6).

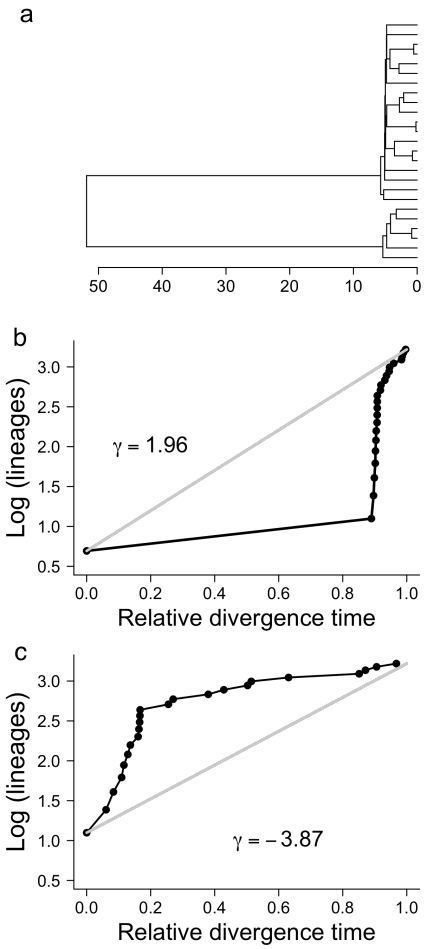

Figure 5.6. Many trees simulated under low ω were characterized by extremely long waiting times between the initial bifurcation and the next speciation event in the phylogenetic tree consisting of extant taxa only. (a) Representative tree from posterior predictive simulations (with ω = 0.1) showing long waiting time following initial bifurcation. Note apparent burst of diversification at approximately 5 mya, which corresponds to the timing of wood-warbler diversification. (b) Lineage accumulation curve for tree shown in (a), with γ much greater than that observed for the wood-warbler data (γ = -3.48). (c) Lineage accumulation curve after omitting the initial bifurcation in the tree; γ for the remaining speciation times is consistent with the wood-warbler dataset. A great number of trees in simulations under ω = 0.1 and ω = 0.5 were characterized by similar long waiting times at the base of the tree, accounting for the bimodal distribution of both γ and tree depth (Fig. 5.5, b). However, when these basal bifurcations are ignored, the distribution of γ is strikingly similar to the wood-warbler data (Fig. 5.5d).

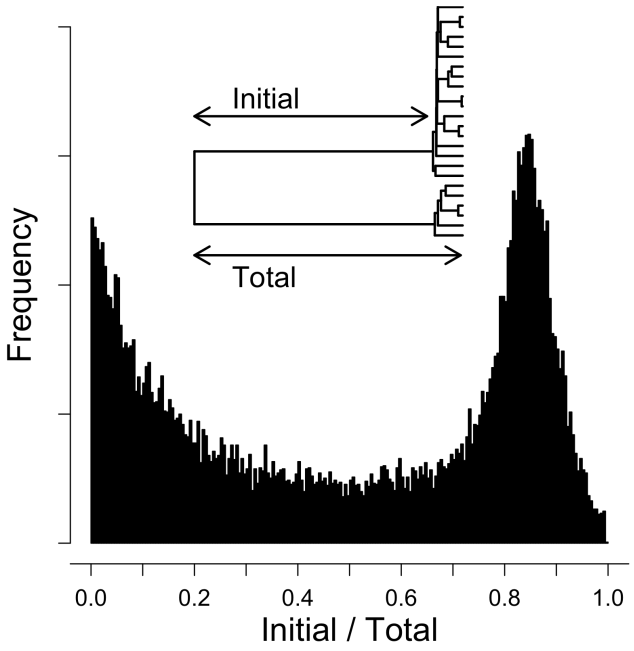

Figure 5.7. Ratio of the initial waiting time to total tree depth for phylogenies simulated under $\omega = 0.1$ model. Many phylogenies failed to eliminate this basal divergence, leading to exceptionally long waiting times until the next speciation event. Trees with high extinction heritability almost always eliminated this long waiting period, such that most trees simulated under $\omega = 1.0$ showed tree depth values consistent with the wood-warbler tree.

Table 5.4. Relative extinction rates (μ / λ) for mean lineage accumulation curves (Fig. 5.4) estimated under the birth-death model assuming either constant or variable speciation and extinction through time.

Tree / model	*Birth-death, constant rate*	*Rate variable I*[2]	*Rate variable II*[3]
Dendroica tree	0	0.106	0
$\omega = 1$	0	0.079	0
$\omega = 0.5$	0	0	0
$\omega = 0.1$	0	0	0

[2] Rate variable I assumes changes in μ and λ occur independently and follow an exponential model. Diversification model is $r(t) = \lambda(t) - \mu(t)$, where $\lambda(t) = \lambda_0 \exp(-kt)$ and $\mu(t) = \mu_0 (1 - \exp(-kt))$. Relative extinction rate for this model is the time-integrated extinction fraction.

[3] Rate variable II assumes both λ and μ decline continuously through time, but that μ is a constant fraction ν of λ (see Materials and Methods)

The HEPT model is implicitly ecological, in that there are limits on clade diversity and that extinction events are immediately followed by speciation events. However, this differs fundamentally from the ecological explanations which have previously been advanced to explain the apparent slowdown in diversification through time observed in many molecular phylogenies (McPeek, 2008; Phillimore and Price, 2008; Rabosky and Lovette, 2008a; Ruber and Zardoya, 2005). Under a model of diversification mediated by ecological opportunity, diversification rates decline through time as the number of species rises within a particular ecological or biogeographic theatre. This may be attributable to a higher frequency of ecological speciation during the early phases of radiations (Gavrilets and Vose, 2005; Rice and Hostert, 1993; Schluter, 2000). Alternatively, if new populations persist better in depauperate environments, then the formation of geographically isolated populations and the effective rate of speciation will decline through time as the number of species increases (Price, 2008). The HEPT model lacks these features and specifies only that speciation events occur immediately after extinction events. Indeed, phylogenies generated under the HEPT model do not even correspond to 'radiations' in any meaningful sense, because rapid lineage accumulation occurs at the base of phylogenetic trees yet is unaccompanied by an increase in species diversity or expansion of ecological space.

A central challenge posed by molecular phylogenetic analyses of diversification has been to explain why estimates of extinction are so often near zero under the birth-death model (Bokma, 2008; Nee, 2006; Rabosky and Lovette, 2008c). This observation is striking in light of the fossil record, which typically suggests that speciation and extinction rates are roughly equivalent when considering sufficiently long time intervals (Alroy, 2000; Alroy, 2008; Pearson, 1996; Sepkoski, 1998). I have shown that, in contrast to the birth-death model, a limited diversity model with zero

net diversification can account for patterns of lineage accumulation observed in real phylogenies.

An alternative to the birth-death model that may also explain the apparent pattern of rapid diversification early is McPeek's (2008) metacommunity model, where he showed that manipulation of ecological similarity between parent and progeny species could generate phylogenetic patterns consistent with the appearance of both temporal declines and increases in diversification through time. In contrast to the model presented here, McPeek's (2008) model predicts that clades showing apparent bursts of lineage diversification early in their history should be characterized by ecological speciation, or at least a tendency for speciation events to be associated with ecological differentiation. Both of these models make predictions that can be tested with additional data, either from the fossil record or patterns of ecological divergence among species within radiations. The HEPT model predicts (i) that diversity should be roughly constant through time; (ii) that pulses of phylogenetically patterned extinction, followed by rapid rebounds in diversity, should be seen in the fossil record; and (iii) that these turnover pulses should correspond in time to the apparent rapid speciation at the base of molecular phylogenetic trees. Clearly there is a need to better integrate perspectives on diversification from molecular phylogenies and fossils.

The HEPT model is best able to explain the wood-warbler data when extinction rates show considerable phylogenetic structure (Fig. 5.5a, b); this is primarily due to the reduced waiting time for coalescence from N = 3 to N=2 species. When extinction is minimally heritable ($\omega = 0.1$), this waiting time may be extremely long relative to the total depth of the tree (Fig. 5.7), and including this 'long stem' leads to old trees showing $\gamma > 0$ (Fig. 5.5c). However, when this initial waiting time is omitted, the waiting times between the remaining speciation events are consistent with

a major decline in the diversification rate through time (Fig. 5.5d; Fig. 5.6). Few studies have explicitly quantified the extent to which extinction in the fossil record is phylogenetically structured (Purvis, 2008), although the sudden disappearance of many groups in their entirety strongly suggests that phylogeny is a strong predictor of extinction (Bininda-Emonds et al., 2007). More evidence emerges from recent work on extinction risk in extant taxa: numerous studies have shown that threatened and endangered species are frequently clustered with respect to phylogeny (Jablonski, 2008; Koh et al., 2004; Purvis et al., 2000; Purvis et al., 2005; Vamosi and Vamosi, 2005; Vamosi and Wilson, 2008). While more work in this area is clearly needed, the HEPT model assumption of phylogenetically structured extinction is likely to be no more problematic than the standard birth-death model, which typically assumes identical extinction rates across all lineages.

One potential weakness of the model is that it appears to require very high speciation and extinction rates to explain the wood-warbler data. The rates given in Table 5.3 (r_2) appear to be at the upper end of the spectrum of previous estimates of speciation (Coyne and Orr, 2004). However, it is somewhat misleading to compare my estimates to previous estimates, for several reasons. First and foremost, previous compilations of speciation rates from phylogenetic data (Coyne and Orr, 2004) have often estimated rates assuming $\mu = 0$. As higher background extinction rates are assumed, the estimated speciation rate will necessarily increase. For example, a simple estimate of the speciation rate based on clade age and species richness (Magallon and Sanderson, 2001) for the wood-warbler tree gives dramatically different results depending on the assumed level of background extinction. If we assume $\mu / \lambda = 0$, the estimated speciation rate is 0.51 lineages/my, but if $\mu / \lambda = 0.99$, the estimate spikes to 4.70 lineages/my. This latter rate is higher than median rates estimated under the HEPT model even when extinction is not phylogenetically structured.

A second point is that many previous estimates of speciation assume homogeneous diversification through time (Magallon and Sanderson, 2001; McCune, 2004; Nee, 2001). Because the HEPT model explicitly estimates rates during pulsed turnover phases, it is only natural that these high rates should be higher than rates averaged across the entirety of evolutionary radiations. The evidence for rate variation through time is substantial (Purvis et al., 2008; Rabosky, 2009), and it does not make sense to compare estimates that assume rate constancy through time to those that do not.

It is important to note that the HEPT model, as formulated here, does not incorporate a mechanism by which phylogenetically patterned extinction might arise. A broad range of biologically-relevant phenomena might lead to apparent heritability of extinction. For example, a key innovation might arise in some particular species that leads to increased species proliferation at the expense of other species. The result might be a turnover process driven by speciation and competitive displacement, where the descendents of the species in which the innovation arose come to dominate the biota. Such a mechanism would likely give patterns consistent with HEPT model predictions.

Limited diversity models with pulsed turnover and phylogenetically structured extinction represent a radically different interpretation of phylogenetic diversification patterns. In contrast to the standard birth-death process, these models can recover apparent bursts of lineage diversification early in a clade's history with high background extinction and zero net diversification. Moreover, because these lineage accumulation patterns can arise despite constant diversity through time, apparent 'bursts' of diversification do not necessarily require the depauperate environments or reduced interspecific competition specified by the ecological opportunity model. These results further imply that early, rapid bursts of lineage accumulation in

molecular phylogenies need not correspond to an adaptive radiation model of diversification, but can arise through temporal variation in turnover rates in conjunction with ecological limits on clade growth. It is too early to determine the extent to which limited diversity models might replace or augment the standard birth-death process, but they have the potential to dramatically change our perspective.

The question is not whether clades radiated; at some point they clearly did. However, this study raises questions about whether lineage accumulation patterns observed in molecular phylogenies correspond in any way to this radiation. We have traditionally focused on identifying key innovations and ecological conditions that promote rapid species accumulation in growing clades, but it may be equally valid to ask whether time-varying turnover rates and constant diversity can account for patterns of speciation in molecular phylogenies.

CHAPTER 6

EXCEPTIONAL AMONG-LINEAGE VARIATION IN DIVERSIFICATION RATES DURING THE RADIATION OF AUSTRALIA'S MOST DIVERSE VERTEBRATE CLADE

Abstract

The disparity in species richness among groups of organisms is one of the most pervasive features of life on earth. A number of studies have addressed this pattern across higher taxa (e.g., 'beetles'), but we know much less about the generality and causal basis of the variation in diversity within evolutionary radiations at lower taxonomic scales. Here we address the causes of variation in species richness among major lineages of Australia's most diverse vertebrate radiation, a clade of at least 232 species of scincid lizards. We use new mitochondrial and nuclear intron DNA sequences to test the extent of diversification rate variation in this group. We present an improved likelihood-based method for estimating per-lineage diversification rates from combined phylogenetic and taxonomic (species richness) data, and we use the method in a hypothesis-testing framework to localize diversification rate shifts on phylogenetic trees. We soundly reject homogeneity of diversification rates among members of this radiation, and we find evidence for a dramatic rate increase in the common ancestor of the genera *Ctenotus* and *Lerista*. Our results suggest that the evolution of traits associated with climate tolerance may have played a role in shaping patterns of diversity in this group.

Introduction

Why do some groups of organisms contain vastly greater numbers of species than other groups? A prominent explanation holds that variation in species richness

among groups can be explained in part by differences in per-lineage rates of species origination and extinction (Slowinksi and Guyer 1989; Mooers and Heard 1997). Many previous studies have tested for variation in diversification rates across large phylogenetic and taxonomic scales, such as angiosperms (Magallon & Sanderson 2001; Sims & McConoway 2003; Davies et al. 2004) or passerine birds (Ricklefs 2006; Phillimore et al. 2006). By comparison, few studies have addressed the extent to which diversification rates vary within species-level radiations, and particularly within groups that do not represent classic adaptive radiations (Ruber & Zardoya 2005; Kozak et al. 2006; McKenna & Farrell 2006). Extending the analysis of diversification rates to species-level radiations of non-model groups is important, because recent work suggests that clade age and not diversification rate might be the dominant signal influencing the distribution of diversity among major animal clades (McPeek & Brown 2007). Understanding the contribution of diversification rate variation to differences in species richness requires more data from evolutionary radiations at lower taxonomic scales.

Australian lizards known as sphenomorphine skinks are a particularly appropriate system in which to test the extent to which diversification rates vary among lineages during continental evolutionary radiations. This group comprises at least 232 species that appear to have diverged *in situ* within mainland Australia (Cogger 2000; Reeder 2003; Skinner 2007), making it by far the most diverse vertebrate radiation in Australia, and one of the most diverse continental radiations across all amniotes. Sphenomorphine skinks occur in virtually all terrestrial habitats in Australia, and body elongation and limb reduction have evolved repeatedly within the group (Greer 1989). Of the fifteen Australian sphenomorphine genera, the genus *Ctenotus* alone includes nearly 100 described species (Cogger 2000). Although it has been suggested previously that *Ctenotus* may have experienced elevated

diversification rates (Reeder 2003), there have been no formal tests for heterogeneity of diversification rates within the sphenomorphine skink radiation.

Here, we present the first molecular phylogenetic analysis of major lineages within *Ctenotus* and use these historical data to test whether it is necessary to invoke among-lineage variation in diversification rates to explain the high disparities in extant diversity among major groups of this radiation. We also test the monophyly of the genus *Ctenotus* to exclude the possibility that the high species diversity of this genus is simply a taxonomic artefact. Finally, we use a likelihood framework based on the birth-death process to test whether high species diversity within *Ctenotus* is the outcome of increased diversification rates within the genus, or whether it reflects a decline in diversification elsewhere in the radiation. Our results indicate a dramatic increase in diversification occurring in the lineage leading to *Ctenotus* and its sister taxon *Lerista* and suggest that arid-adapted lineages may have diversified explosively in response to the expansion of the Australian arid zone over the past 20 million years.

Materials and Methods

Taxon Sampling

We obtained tissue samples for 34 species of *Ctenotus*, including multiple representatives of the twelve recognized 'species groups' (Storr et al. 1999) and most of the phenotypic diversity present in this high-diversity genus (Greer 1989; Cogger 2000). A summary of major morphotypes within *Ctenotus* is given as supplemental material (Table S1 in Appendix I). Genomic DNA was extracted using DNeasy Tissue kits (Qiagen). We sequenced 718 bp of mitochondrial ND4 and 181 bp of three flanking tRNAs (tRNA-ser, tRNA-his, tRNA-leu) using primers and protocols described in Reeder (2003). In addition, we sequenced 671 bp of the nuclear ATP synthetase β-subunit intron, after Skinner (2007). We aligned new DNA sequences

for *Ctenotus* to those used in a previous analysis of sphenomorphine skink relationships (Reeder 2003; Skinner 2007). The full data matrix contained 81 species, including representatives of all Australian sphenomorphine genera, and six outgroup taxa. In addition to ND4, tRNA, and ATP synthetase intron sequences, we included 1442 bp of mitochondrial 12S and 16S rRNA sequences for 48 of these species that were available from previous studies (Reeder 2003; Skinner 2007). Alignment of ND4 did not require gap insertion and was straightforward. Ribosomal 12S and 16S alignments used secondary structure modes from the European ribosomal database (Wuyts et al. 2001), tRNAs were aligned after Macey & Verma (1997), and ATP synthetase intron sequences were aligned using Clustal X (Thompson et al. 1997) with default parameters. Genbank accession numbers, museum voucher numbers, and sampling localities for all sequences included in this study are available in Appendix I (table s2).

Phylogenetic analysis

Bayesian phylogenetic analyses were performed on the combined dataset using separate partitions for intron (ATP synthetase), protein coding (ND4), and structural (tRNAs, rRNAs). Appropriate models of molecular evolution for the three partitions were determined using MrModeltest v2.2 (Nylander 2004). Recognition of additional data partitions (Brandley et al. 2005) did not change results, and we present analyses based on these three partitions only. We performed two simultaneous runs of Metropolis-coupled Markov chain Monte Carlo (MCMCMC) sampling using the parallel processor version of MrBayes v3.1.2 (Altekar et al. 2004), with four incrementally heated chains run simultaneously for 10 million generations and trees sampled every 1000th generation. We discarded the first 7.5 million generations as burn-in; we checked for convergence by (i) testing for stationarity of logL values, (ii)

by plotting the posterior probabilities for nodes as a function of the number of generations, and (iii) examining standard deviations of split frequencies for independent runs (Huelsenbeck et al. 2001). We performed two additional runs using identical conditions to verify that results converged on the same posterior distribution of trees

Variation in diversification rates

We obtained an ultrametric tree for diversification analyses by applying penalized likelihood (PL; Sanderson 2002) to the Bayesian consensus phylogram using the software package r8s (Sanderson 2003). Prior to rate smoothing, we removed all non-Australian sphenomorphine taxa from the tree, and applied cross-validation procedure implemented in r8s with an additive penalty to select an optimal smoothing parameter. The diversification rate analyses discussed below are robust to absolute age estimates and require only relative divergence times. However, to facilitate comparison with other Australian radiations for which approximate timescales are available (e.g., Jennings et al. 2003; Crisp et al. 2004), we inferred the age of the basal split among Australian sphenomorphine species. No fossil calibration points exist within either the sphenomorphine radiation or closely related outgroup taxa (Martin et al. 2004), so we used published rates of squamate mtDNA sequence evolution (Zamudio & Greene 1997; Calsbeek et al. 2003; Richmond & Jokusch 2007). We determined mean pairwise maximum likelihood (ML) distances between taxa in the two basal clades under the best fit model and parameters inferred by MrModeltest for this data partition and estimated the timing of the split under an mtDNA divergence rate of 0.895 % corrected sequence divergence per million years (range: 0.47 – 1.32 %; Zamudio & Greene 1997).

Diversification rate analyses were conducted on the genus-level tree created by pruning all but one lineage per genus from the PL tree; species totals were assigned to each terminal following Wilson & Swan (2003). We tested whether some groups of Australian skinks are characterized by exceptionally high or low diversification rates using the approach of Magallon & Sanderson (2001). We plotted the number of species within Australian sphenomorphine genera as a function of stem clade age, as inferred from the PL chronogram. The stem clade age is simply the age of the divergence between the crown group and its extant sister group, corresponding in this case to genus age; this assumes that genera are monophyletic. Because we do not have complete sampling of all sphenomorphine species, we cannot distinguish between crown-group and stem-group ages, and these divergence times thus represent the transition time from a single ancestral species to a clade with some level of standing diversity in the present.

We then considered the Australian sphenomorphine radiation as a whole and estimated the net diversification rate, r, or the difference between the speciation rate (λ) and the extinction rate (μ). We compared the extant diversity of sphenomorphine genera to the 95% confidence intervals around expected species diversity for clades of similar age that have diversified with the net diversification rate observed for the overall radiation (Magallon & Sanderson 2001, eqn 10).

Net diversification rates were estimated for the entire sphenomorphine radiation using two methods. First, we used the whole clade estimator described in Magallon and Sanderson (2001; eqn 7). This estimator is a function of three parameters: the number of species in the clade, the age of the clade, and the ratio of the extinction rate to the speciation rate, μ / λ (denoted by a). Even with a constant net diversification rate, the expected outcome of a stochastic diversification process can vary with the extinction fraction a (e.g., Raup 1985). We therefore computed

confidence intervals around expected diversity under $a = 0$ and $a = 0.99$; these values represent extremes on a continuum of relative extinction rates. We computed one-tailed probabilities of the observed diversity of each clade, given the estimated overall diversification rate and clade age, to test whether sphenomorphine skinks are characterized by an excess of species-rich or species-poor lineages; we combined individual p-values using the Z-transform method (Whitlock 2005; additional information Appendix I, Table S4).

We repeated these analyses using an estimator of r that uses both taxonomic and phylogenetic data, because our data consist of terminal taxa of known diversity (e.g., numbers of species within genera) as well as the phylogenetic relationships of these taxa. Our estimator generalizes the Yule process estimator of Sanderson & Wojciechowski (1996) to the birth-death process. The taxonomic data consists of T terminals in a molecular phylogeny with branch length data. The length of each terminal branch represents the stem-clade age of each taxon, or the maximum elapsed time between a single ancestral species and n descendant species. Let λ and μ represent per lineage probabilities of speciation and extinction, and define $r = \lambda - \mu$ and $a = \mu / \lambda$. Under the birth-death process, the probability of observing n species in the present given a single ancestral species is given by (Kendall 1948)

$$\Pr(n \mid t, r, a) = (1 - \alpha)(1 - \beta)\beta^{n-1} \tag{6.1}$$

where

$$\beta = \frac{\exp(rt) - 1}{\exp(rt) - a} \tag{6.2}$$

and $\alpha = a\beta$,

and the probability of observing zero descendant species is

$$\Pr(n = 0 \mid t, r, a) = \alpha \tag{6.3}$$

This is identical to Raup (1985, eqn A17) for the case of a single ancestral species. A maximum likelihood estimator of *r* based on the taxonomic data only can be obtained from the product of the probabilities of the taxonomic data:

$$L_T = \prod_{i=1}^{T} \Pr(n_i \mid t_i, r, a) \tag{6.4}$$

This is the likelihood used in Paradis (2003). However, if a diversification process has left zero descendants in the present, it cannot be observed. Thus, we condition Pr ($n \mid t$) on the probability that the clade has survived to the present (Nee et al. 1994), which it does with a probability of $1 - \alpha$. After conditioning,

$$\Pr(n \mid t, r, a, n \geq 1) = (1 - \beta)\beta^{n-1} \tag{6.5}$$

Multiplying the probabilities for each of the *T* terminals together and taking the log gives the log likelihood of the taxonomic data, or

$$\log L_T = \sum_{i=1}^{T} \log(1 - \beta_i) + \sum_{i=1}^{T} (n_i - 1) \log \beta_i \ . \tag{6.6}$$

The phylogenetic component consists of *N* internal branches, each beginning with a single lineage and ending with a speciation event. The probability density of each branching event, conditional upon each lineage surviving to the present, is given in Nee et al. (1994, eqn 17). For the special case of a single ancestral lineage, we have

(6.7) $$\Pr(t \mid a, r) = r\left[1 - a\exp(-rx)\right]^{-1}\exp(-rt)$$

where x is the birth time or branching time (Nee et al. 1994; Nee 2001) of the lineage in time units before present. By multiplying the probability densities for each of the N branching events and taking the log of both sides, we obtain the log likelihood that each of the N internal branches in the phylogeny will result in exactly two descendant species after some amount of time

(6.8) $$\log L_P = N\log(r) - r\sum_{i=1}^{N} t_i - \sum_{i=1}^{N}\log\left[1 - a\exp(-rx_i)\right]$$

where n_i is the diversity of the i'th terminal, t_i is the stem age of the terminal, and a is the extinction fraction. This is the likelihood of the phylogenetic data (Lp) and differs from Paradis (2003), in that Paradis (2003, eqn 2.5) includes the likelihood that each of the T terminal lineages does not undergo a speciation event on the terminal branch. This is the case for a complete phylogeny, but we know that at least some of the T terminals have left more than one progeny in the present; furthermore, the likelihoods of the terminal branches have already been considered in the taxonomic portion of the expression (L_T; eqn 6.6).

The log-likelihood of the combined taxonomic and phylogenetic data is simply $\log L = \log L_P + \log L_T$ and can be found under any assumed level of background extinction. Maximum likelihood estimates of r were found using the one-dimensional 'optimize' routine in R (http://cran.r-project.org/).

Shifts in diversification rate

We tested for shifts in diversification rate within the Australian sphenomorphine radiation by contrasting the likelihood of the data under a model with equal diversification rates for all lineages to the likelihood under a model where an ancestral diversification rate r_1 shifts to a new rate r_2 along some branch in the tree (Sanderson 1994; Sanderson and Wojciechowski 1996). To compute likelihoods and to estimate *r*, we used the birth-death estimator based on phylogenetic and taxonomic data described in this study. For this model, which we refer to as the flexible-rate model, we sequentially split the tree at each branch and optimized *r* onto the resulting pair of subtrees. The node resulting in the maximum combined likelihood for the bipartite tree was the maximum likelihood estimate of the shift point. Analyses were conducted under extinction fractions of $a = 0$ and $a = 0.99$.

We predicted that the clade containing *Ctenotus* and *Lerista* would show an increase in diversification rates relative to other lineages simply because of the greater diversity of these genera relative to other sphenomorphine genera. However, an alternative hypothesis is that this clade retained an ancestral but elevated diversification rate, while another clade or clades exhibited a decline in diversification. To distinguish between these possibilities, we repeated our flexible-rate analysis with the constraint that the subtree partition containing *Ctenotus* must include the root node. Thus, no rate shift was permitted along the path between the root node and *Ctenotus*. If the likelihood of the data under this constrained model is similar to that under the flexible-rate model, we cannot distinguish between an increase in diversification rates for *Ctenotus* and a decrease in diversification rates for some other clade. To avoid conditioning results on any particular tree topology and branch lengths, we sampled 1000 phylogenies from the set of post-burnin topologies (1 tree

every 5,000 generations) and applied PL to each. We repeated our analyses on this set of trees to generate the posterior distribution of likelihood differences between these two variable-rate diversification models, with the expectation that the rate-decrease model would consistently provide a poorer fit to the data than the flexible-rate model. Functions for all diversification analyses described in this study have been included in the LASER library for the R programming language (Rabosky 2006a).

Results and discussion

Phylogenetic relationships

Bayesian phylogenetic analysis supports *Ctenotus* as monophyletic with a high level of confidence (Figure 6.1), implying that the exceptional species diversity within *Ctenotus* cannot be explained as a taxonomic artefact. In addition, our results strongly support sister clade relationship between *Ctenotus* and *Lerista*, the other highly diverse genus of sphenomorphine skink (79 species). Although taxon sampling of *Lerista* is limited, morphological and molecular data (Greer 1989; Skinner, pers. comm.) strongly suggest that the genus is monophyletic.

Our results corroborate those of several previous studies (Reeder 2003; Skinner 2007) and provide further support for the monophyly of Australian sphenomorphine skinks, as well as a basal split between *Notoscincus* and all other Australian sphenomorphine lineages. These results were also supported by parsimony and likelihood bootstrap analyses (Appendix I, Table S3). As found previously (Reeder 2003), several sphenomorphine genera are not monophyletic (e.g., *Eulamprus*, *Glaphyromorphus*). Taken together, monophyly of *Ctenotus* and of *Ctenotus* plus *Lerista* suggests higher diversification rates for these lineages relative to the remainder of the Australian sphenomorphine radiation, because fully 75% of the species in the radiation fall within these genera.

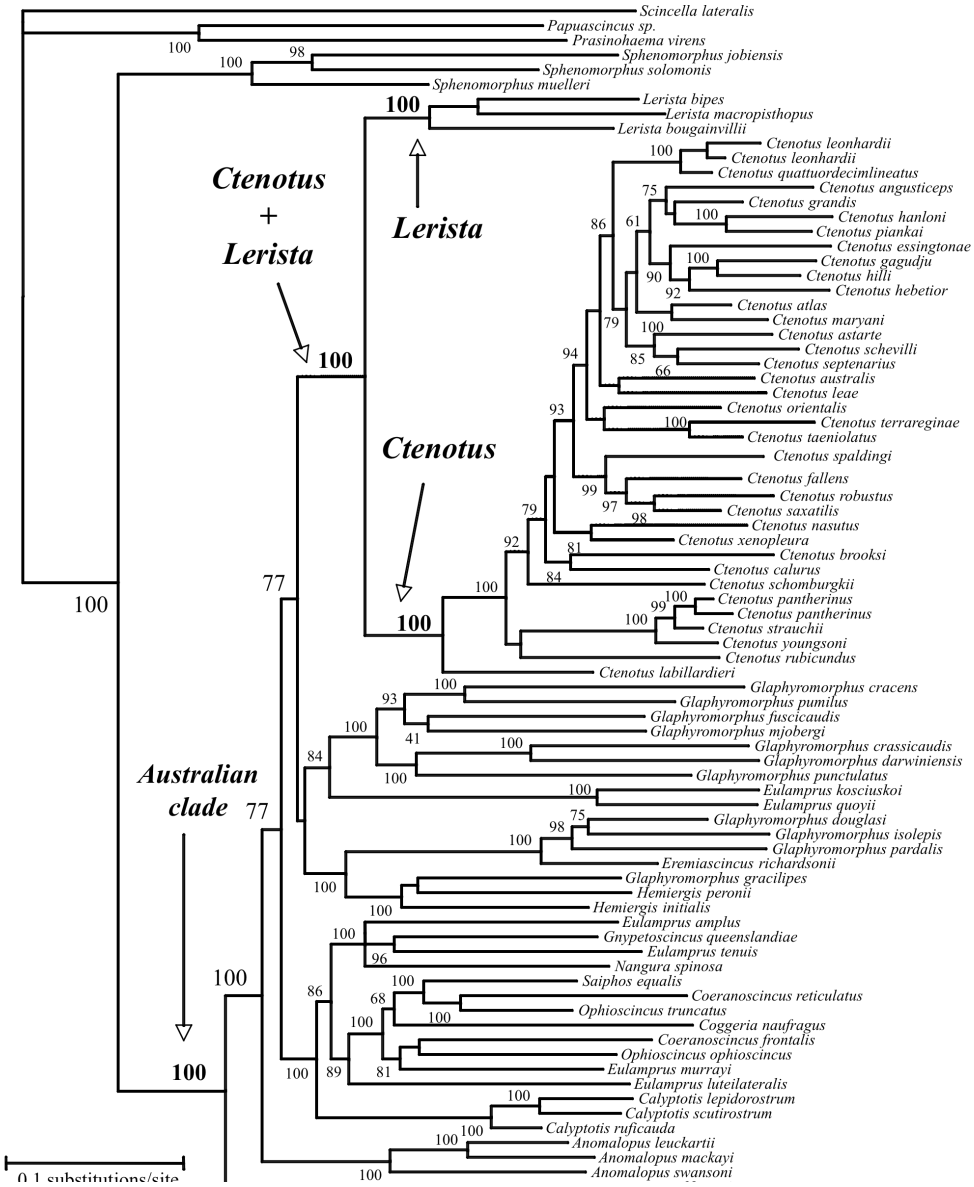

Figure 6.1. Bayesian consensus phylogram for 75 Australian and 6 non-Australian sphenomorphine skink species based on combined analysis of mtDNA and nuclear intron sequences. Only posterior probabilities (x 100) greater than or equal to 60 are shown. Branch lengths are means from the posterior distribution of sampled trees.

Variation in diversification rates

Cross-validation analysis of penalized likelihood (PL) trees with different smoothing parameters and of a tree with branch lengths estimated under a molecular clock constraint indicated optimal performance of PL with a smoothing parameter of 100. We pruned the PL tree to include 23 lineages for which we could assign levels of taxonomic diversity (Figure 6.2). We inferred a basal divergence time of 28.2 million years before present (mya; range: 19.1 – 53.6 mya) using ML distances between *Notoscincus* and other Australian sphenomorphine skinks. In the absence of robust fossil calibration points, we prefer to treat this age estimate provisionally, but note that a recently published substitution rate for squamate reptile ND4 (1.6 % Myr $^{-1}$) would imply that the radiation has occurred even more recently (Feldman and Spicer 2006). With the exception of the absolute magnitude of inferred diversification rates, all results are independent of the proposed age of the sphenomorphine radiation.

We were able to assign taxonomic diversity levels to the four paraphyletic genera, because taxon sampling included most or all species within three of those genera (2/2 *Coeranoscincus*, 2/3 *Ophioscincus*, 11/13 *Glaphyromorphus*), and because major lineages within *Eulamprus* correspond to phenotypically and ecologically distinctive categories (O'Connor and Moritz 2003). To verify that results were robust to alternative species richness assignments, we generated 100 datasets by randomly partitioning generic diversity into paraphyletic lineages (e.g., the 15 recognized species of *Eulamprus* were divided at random among *Eulamprus* lineages *A*, *B*, *C*, *D*, and *E* in each replicate dataset) and repeated the full complement of diversification analyses on each. Shuffling diversity levels within paraphyletic genera in this fashion did not change results described below (Appendix II, Figure S1).

Comparisons of observed species diversity within terminal taxa relative to those expected under a homogeneous net diversification rate across all Australian

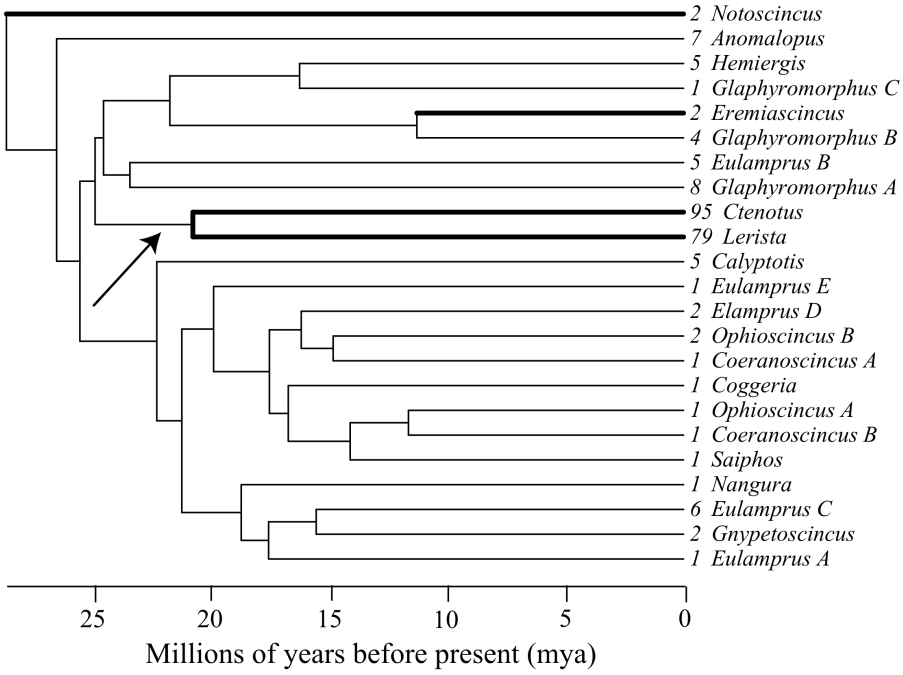

Figure 6.2. Penalized likelihood chronogram for Australian sphenomorphine diversification derived from Bayesian consensus phylogram, indicating taxonomic diversity of each terminal. Basal divergence of 28.2 mya is based on estimated mtDNA rate of 0.895 substitutions Myr^{-1}. The four paraphyletic genera (*Coeranoscincus*, *Eulamprus*, *Glaphyromorphus*, *Ophioscincus*) were assigned taxonomic diversity levels based on the full phylogeny (fig. 1), or from prior information regarding morphological species groups. Results of diversification analyses are robust to species richness assigned to paraphyletic genera (Appendix I, fig. s1). Maximum likelihood estimate of rate-shift location inferred under two-rate diversification model is indicated by arrow. Bold taxon names and branch lengths denote the four arid-adapted taxa.

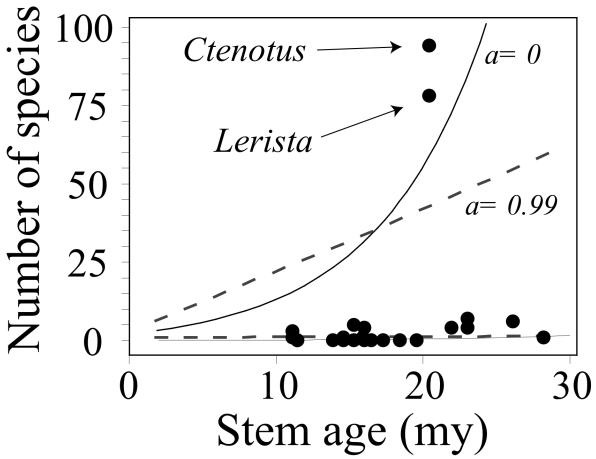

Figure 6.3. Relationship between stem clade age and extant diversity for Australian sphenomorphine skinks. Lines indicate the 95% confidence envelope around the expected species diversity through time for a clade that diversifies with a rate equal to that observed across the Australian sphenomorphines considered as a whole under relative extinction rates of $a = 0$ (solid line) and $a = 0.99$ (dashed line). *Ctenotus* and *Lerista* have an excess of species and the remaining lineages have a deficit of species assuming a constant diversification rate for the entire sphenomorphine clade ($a = 0$, $p < 0.013$; $a = 0.99$, $p < 0.032$). Net diversification rate was estimated from combined taxonomic/phylogenetic data.

Table 6.1. Model-based analysis of diversification rates in Australian sphenomorphine skinks. (Flexible-rate model considers all possible bipartitions of tree and finds bipartition giving highest likelihood when net diversification rates are optimized separately to each partition. Rate-decrease model assumes that *Ctenotus* and *Lerista* have retained the ancestral diversification rate present at the root node).

Model	*Constant rate*	*Flexible-rate*	*Rate-decrease*
	LogL (ΔAIC[2])	LogL (ΔAIC)	LogL (ΔAIC)
a = 0	-117.6 (19.0)[3]	-106.1 (0)	-111.6 (10.9)
parameters[1]	r = 0.135	r = 0.086; r_{CL} = 0.218	r = 0.073; r_{CL} = 0.160
a =0.99	-136.3 (32.7)[3]	-117.9 (0)	- 124.7 (13.6)
parameters[1]	r = 0.005	r = 0.002; r_{CL} = 0.030	r = 0.0003; r_{CL} = 0.005

[1] Maximum likelihood estimate of the net diversification rate, *r*, in lineages Myr $^{-1}$. r_{CL} is the net diversification rate of the subtree partition containing *Ctenotus* and *Lerista*.

[2] ΔAIC is the difference in AIC scores between each model and the overall best-fit model

[3] The data reject the constant rate model in favour of the flexible-rate model under a = 0 and a = 0.99 ($\chi^2 \geq 23$; $p < 0.001$). Rate-decrease model does not show a simple nested relationship with flexible-rate model, but AIC strongly favours the flexible-rate model.

sphenomorphine skinks reject the null hypothesis that net diversification rates have been constant among lineages ($p < 0.013$, $a = 0$; $p < 0.031$, $a = 0.99$; Figure 6.3).

Regardless of whether sphenomorphine diversification rates are estimated using the combined taxonomic/phylogenetic approach described in this study (Figure 6.3) or the whole-clade estimator ($p < 0.001$; Appendix II, Table S4), there is clearly an excess of species-poor or species-rich lineages. Although this test is conservative (Appendix II, Table S4e), these results are robust to assumptions about the underlying model of extinction. Therefore, clade age alone does not explain the striking disparity in species richness among sphenomorphine skinks.

Shifts in diversification rate

The data clearly reject the constant diversification model in favour of the flexible-rate model ($p < 0.001$; Table 6.1). The ML estimate of the shift point in the flexible-rate model is the node corresponding to the most recent common ancestor (MRCA) of *Ctenotus* and *Lerista* (Figure 6.2) under extinction fractions of both $a = 0$ and $a = 0.99$; ML estimates of r for *Ctenotus*/*Lerista* and for all other Australian sphenomorphine lineages suggest that the net diversification rate within this clade has increased approximately 2.5-fold ($a = 0$) to 15-fold ($a = 0.99$) relative to that observed across the remainder of the tree. While palaeontological evidence supports the view that relative extinction rates have generally been high (e.g., Stanley 1979; Gilinsky 1994; Alroy 1996), we are aware of no evidence suggesting that extant clades have diversified in the absence of extinction. It is thus likely that the true magnitude of the rate increase for *Ctenotus* and *Lerista* exceeded that inferred under $a = 0$. Rate shifts at other internal nodes are far less likely than at the MRCA of *Ctenotus*/*Lerista*, as assessed by the difference in likelihood scores (ΔL) between the best-fit location of the rate shift and alternative nodes (figure 6.2): $MRCA_{Ctenotus\text{-}Hemiergis}$, $\Delta L = 4.06$;

MRCA$_{Ctenotus\text{-}Anomalopus}$, $\Delta L = 9.99$; shift on *Ctenotus* branch only, $\Delta L = 7.05$; shift on *Lerista* branch only, $\Delta L = 8.36$ (ΔL given for $a = 0$, but results for $a = 0.99$ are qualitatively similar).

The hypothesis that *Lerista*/*Ctenotus* have merely retained high diversification rates from the ancestral sphenomorphine lineage is not supported, as this scenario would require that diversification rates have decreased multiple times throughout the tree. The model specifying increased diversification at the *Lerista*/*Ctenotus* MRCA performs much better than a model where no rate increase is permitted between the root node and *Lerista*/*Ctenotus* (rate-decrease model; table 6.1), as assessed by the AIC. This result is not conditional on the topology and branch lengths shown in Figure 6.2, because the posterior distribution of the difference in likelihood scores between the flexible-rate model and the rate-decrease model strongly favours the flexible-rate model (Appendix I, Figure S2).

Because our analyses are partially dependent on taxonomy, our results could have been influenced by a failure to adequately account for true species diversity within genera. Very few intraspecific phylogeographic studies have been conducted on sphenomorphine skinks, and we predict that future investigations will reveal morphologically cryptic species within many taxa. However, this issue would only influence our results if *Lerista* and *Ctenotus* harbour proportionately fewer cryptic species than other taxa.

Aridification of Australia and diversification rates

The progressive expansion of the Australian arid zone during the past 25 million years appears to have catalyzed increased diversification in a number of sclerophyllous plant clades, including eucalypts and casuarinas (Crisp et al. 2004). Likewise, some lineages of agamid lizards experienced rapid diversification during the

aridification of Australia (Harmon et al. 2003; Hugall & Lee 2004; Rabosky 2006b). The majority of species within both *Lerista* and *Ctenotus* occur in the arid and semi-arid regions of Australia, suggesting a possible link between aridification and diversification rates in this group. Temporal calibration of the sphenormophine tree (Figure 6.2) yields dates for the radiations of *Ctenotus* and *Lerista* that accord well with the known chronology for the aridification of Australia and that show broad congruence with the timing of diversification in these other arid-adapted groups.

Given the diversity of these two genera in the Australian arid zone, it is striking that only a few other lineages of sphenomorphine skinks (Figure 6.2) can be considered arid zone taxa – two species each of *Notoscincus* and *Eremiascincus*. Thus, while nearly half of all species of sphenomorphine skinks occur in the arid and semi-arid regions of the continent, this diversity is concentrated within just 4 of 23 major lineages (Figure 6.2). Yet the arid and semi-arid climatic regions of Australia are defined by a climatic zone that occupies more than three-fourths of the continent's surface area (James & Shine 2000). That so few major lineages have succeed in entering a geographically vast region that nonetheless comprises a small fraction of the total climatic diversity in Australia (James & Shine 2000) suggests the possibility that phylogenetic conservatism of climatic tolerances (Ricklefs & Latham 1992; Wiens & Graham 2005) may underlie the exceptional diversification of a few sphenomorphine lineages. Such phylogenetic conservatism of traits is hypothesized to limit the dispersal of lineages between tropical and temperate regions (Ricklefs & Latham 1992; Wiens et al. 2006) and among elevational zones in the montane tropics (Wiens et al. 2007); we similarly hypothesize that such constraints have influenced the ability of lineages to shift from mesic to arid environments.

Can these diversification patterns be explained by niche conservatism in conjunction with the larger area of the arid zone, relative to the more mesic regions of

continental Australia? There is considerable evidence that geographic area occupied by a clade exerts a strong effect on diversification rates (Losos and Schluter 2000; Ricklefs 2003; Davies et al. 2005). The mesic-to-arid gradient might act as an environmental filter, limiting the dispersal of lineages between those environments, and the larger area of the arid zone would thus provide an expanded theatre of diversification for clades that successfully made this transition. If arid Australia harbours a diversity of sphenomorphine species proportional to its geographic area, and if a restricted number of clades have been able to enter this climatic zone, we would expect to observe increased diversification rates in those clades as a consequence of the interaction between geographic area and phylogenetic conservatism of climatic tolerance. Indeed, *Ctenotus* species richness within major climatic regions of Australia is proportional to the geographic area occupied by those zones (James & Shine 2000). This suggests that part of the answer to 'why do so many *Ctenotus* species occur in the arid zone' (Pianka 1972; Morton & James 1988; James & Shine 2000) may come from understanding why other sphenomorphine clades fail to show this pattern.

The phylogenetic conservatism/geographic area hypothesis might only be part of the explanation for the increased diversification of *Ctenotus* and *Lerista* relative to other sphenomorphine skinks. Assuming that the ancestral Australian sphenomorphine was not arid adapted, as is likely based on the general climatic tolerances of sphenomorphine outroup taxa, only four major lineages have successfully made the transition to arid environments (if the putative shift in climatic tolerance occurred in the ancestor of *Ctenotus* and *Lerista*, then only three lineages have made this transition). Yet two of these lineages failed to radiate (*Notoscincus* and *Eremiascincus*). Previous studies have found that the *Ctenotus* lineage is characterized by a substantial increase in both critical thermal maximum temperatures

and preferred active temperatures relative to other sphenomorphine species, including *Eremiascincus* (Benett & Alder-John 1986; Huey & Bennett 1987; Garland et al. 1991). Although data are needed for *Lerista* and *Notoscincus*, these results suggest the possibility that traits related to thermal physiology might underlie the dramatic radiation of these groups in the arid zone. However, despite a close phylogenetic relationship, *Ctenotus* and *Lerista* have followed radically different evolutionary paths: *Ctenotus* is diurnal, surface active, and shows minimal interspecific variation in body shape, whereas *Lerista* is fossorial to cryptozoic, frequently nocturnal, and shows a tremendous range of interspecific variation in limb reduction and body elongation (Greer 1989).

Conclusions

This study demonstrates exceptional heterogeneity in diversification rates within a major continental radiation. The high species richness of *Ctenotus* and *Lerista* relative to other genera of Australian skinks cannot be explained by clade age or as a taxonomic artefact. Tremendous variation in diversification rates appears to have characterized several continental plant radiations (Hodges & Arnold 1995; Klak et al. 2003; Hughes & Eastwood 2006), but we are unaware of a comparable, sustained contrast in diversification rates within animal radiations restricted to a geographically contiguous region. The diversification of Australian sphenomorphine skinks has occurred against a background of apparently strong phylogenetic niche conservatism, and we speculate that a key physiological innovation may have catalyzed diversification in some, but not all, arid-adapted lineages.

CHAPTER 7

DENSITY-DEPENDENT DIVERSIFICATION IN NORTH AMERICAN WOOD-WARBLERS

Abstract

Evidence from both molecular phylogenies and the fossil record suggests that rates of species diversification often decline through time during evolutionary radiations. One proposed explanation for this pattern is ecological opportunity, whereby an initial abundance of resources and lack of potential competitors facilitates rapid diversification. This model predicts density-dependent declines in diversification rates, but has not been formally tested in any species-level radiation. Here we develop a new conceptual framework that distinguishes density-dependence from alternative processes that also produce temporally declining diversification, and we demonstrate this approach using a new phylogeny of North American wood-warblers. We show that explosive lineage accumulation early in the history of this avian radiation is best explained by a density-dependent diversification process. Our results suggest that the tempo of wood-warbler diversification was mediated by ecological interactions among species and that lineage and ecological diversification in this group are coupled, as predicted under the ecological opportunity model.

Introduction

One of the most striking features of evolutionary radiations is a tendency for species level diversification rates to decline through time. This pattern has long been recognized in the fossil record, where explosive but transient bursts of diversification appear to follow both mass extinctions (Sepkoski 1998) and the invasion of previously unoccupied adaptive zones (Simpson 1953; Stanley 1973). A large number of studies

have used molecular phylogenies of extant taxa to document a pattern of early, rapid diversification, followed by temporal declines in diversification rates (e.g., Lovette & Bermingham 1999; Harmon et al. 2003; Ruber & Zardoya 2005; Weir 2006; Rawlings et al. 2008; Phillimore & Price 2008). Although several potential biases can generate spurious shifts in diversification rates inferred from molecular phylogenies (Nee 2001; Revell et al. 2005), methodological improvements (Pybus & Harvey 2000; Rabosky 2006a) continue to support the phenomenon of declining diversification rates through time in species-level radiations.

One potential biological explanation for this pervasive pattern is that evolutionary radiations are facilitated by ecological opportunity (Schluter 2000), whereby speciation is most likely when resources are abundant and potential competitors scarce. As a radiation progresses, ecological 'niche space' becomes increasingly saturated, resulting in fewer opportunities for speciation (Walker & Valentine 1984; Valentine 1985). Under such a model, speciation rates are predicted to show density-dependence (Nee et al. 1992), because the rise in species diversity through time would be mirrored by corresponding decline in the speciation rate. In a meta-analysis of 45 avian radiations, Phillimore & Price (2008) found widespread evidence for temporal slowdowns in diversification rates and speculated that ecological limits on clade growth resulted in density-dependent speciation. However, there have been no formal tests of density-dependent diversification in any evolutionary radiation, due to lack of an appropriate statistical framework. Although methods are available for detecting temporal declines in diversification rates (e.g., Pybus & Harvey 2000; Rabosky 2006b), it has not been possible to discriminate between density-dependence and other processes that might also result in temporal declines in diversification rates.

Here we develop a novel conceptual framework for testing whether diversification rates show density-dependence, and we explore the role of this process during the radiation of continental North American wood-warblers (Parulidae) in the speciose genus *Dendroica*. *Dendroica* warblers are an ecologically appropriate group in which to test for density dependence in diversification, as this process is most likely to be driven by interspecific competitive interactions. *Dendroica* species diversity is high in many local North American assemblages, but the composition of those assemblages is variable among sites (Lovette and Hochachka 2006). The matrix of potential species interactions is even more complex when integrated across the history of this group, as most *Dendroica* species have persisted through climate cycles that would have further scrambled their geographic ranges and spatial associations. These warblers are a classic example of behavioral niche differentiation, in which co-occurring species differ in subtle aspects of their foraging and breeding behavior (MacArthur 1958; Price et al. 2000). Our previous studies based on a time-calibrated mitochondrial DNA (mtDNA) phylogeny found that the *Dendroica* group underwent an explosive burst of diversification early in its history, followed by a pronounced decline in the rate of lineage accumulation (Lovette & Bermingham 1999). Taken together, these observations suggest the possibility that broad-scale patterns of diversification in the *Dendroica* warblers might be related to ecological interactions among species across evolutionary timescales.

Our statistical approach extends the birth-death model that has been used previously for inference on diversification rates (Nee et al. 1994; Barraclough & Vogler 2002; Rabosky 2006b; Nee 2006) to speciation rates that vary continuously through time. We apply the method to an improved phylogeny of *Dendroica* warblers that is complete at the species level and which is based on both mtDNA and nuclear sequence loci. Our results indicate that the observed pattern of speciation in *Dendroica*

is best approximated by a density-dependent diversification process and suggest that ecological interactions among species can leave an imprint on evolutionary history that can be reconstructed from molecular phylogenies alone.

Materials and Methods

Taxon sampling and phylogenetic analyses

We reconstructed relationships among all species in the *Dendroica* radiation, including four species traditionally assigned to the genera *Parula*, *Setophaga*, and *Wilsonia* that fall within this well-supported clade (Lovette & Bermingham 2002). We sequenced a total of six mitochondrial protein-coding genes and six nuclear intron loci. All genes and loci were obtained from all taxa, except for *Dendroica chrysoparia*, for which we had only older museum skin material, and hence from which we obtained only the mitochondrial-encoded ND2 sequence. Reconstructions were based on a total of 5261 nucleotides of protein-coding mtDNA sequence and 4296 aligned nucleotides of nuclear intron sequence. Although the phylogenetic reconstructions included all taxa within the *Dendroica* radiation, our subsequent tests of diversification rates and their potential density dependence employed a pruned tree that excluded lineages restricted to West Indian islands, as those isolated island taxa could not have been involved in interspecific interactions in continental warbler communities.

We used BEAST v1.4.6 to simultaneously infer topologies and relative divergence times under a relaxed-clock model of sequence evolution (Drummond et al. 2006; Drummond & Rambaut 2007). Because the diversification rate analyses described below require only that trees be calibrated to relative timescales, we used a model of uncorrelated but lognormally-distributed substitution rates and fixed the mean rate at 1.0. We recognized four partitions with independent evolutionary

parameters (three mitochondrial codon positions plus nuclear DNA) and assigned a GTR+G+I model of sequence evolution to each, following analyses with MrModeltest v2.2 (Nylander 2004). We performed 11 runs of Markov Chain Monte Carlo (MCMC) on the combined dataset, sampling parameters every 30,000 generations and discarding the first 5 million generations from each as burnin. Post-burnin parameters and trees were combined across runs for a total of 341 million generations of MCMC sampling. We placed uniform [0, 100] priors on parameters of the substitution rate matrix and default priors on all other parameters. We assessed convergence on the posterior distribution by calculating effective sample sizes for evolutionary parameters using Tracer v1.4 (Drummond et al. 2006).

Models and parameter estimation

To test for density-dependent diversification, one cannot simply contrast the likelihood of phylogenetic data under a density-dependent diversification model to the corresponding likelihood under a constant-rate model of diversification. There are many reasons why diversification rates might appear to decline gradually through time that have nothing to do with density-dependent cladogenesis. For example, both incomplete taxon sampling and artefacts of phylogeny reconstruction can result in spurious declines in diversification rates through time, even when rates have not changed (Nee et al. 1994; Nee 2001; Revell et al. 2005). It may be the case that density-dependent models fit such data better than constant-rate models, simply because these models provide a crude approximation of continuous declines in diversification rates. In Appendix II (Figure S1), we show that a constant rate diversification process in conjunction with incomplete taxon sampling can strongly mimic density-dependent diversification when the data are simply fitted with density-dependent and constant rate diversification models. We argue that the relevant null

hypothesis for density-dependent diversification is a model where diversification rates are permitted to vary continuously through time but not directly as a function of the number of lineages in existence.

We considered both exponential and linear models of density-dependent diversification (Nee et al. 1992; Rabosky 2006a). Under the exponential model, the speciation rate λ is modelled as

$$\lambda(t) = \lambda_0 N_t^{-x} \tag{7.1}$$

where λ_0 is the initial speciation rate, N_t is the number of lineages in existence at time t in a reconstructed phylogeny, and x determines the magnitude of the rate change as a function of N_t. Note that x = 0 implies constant speciation through time. For the linear model,

$$\lambda(t) = \lambda_0\left(1 - {N_t}/{K}\right) \quad , \tag{7.2}$$

where K is analogous to the carrying capacity parameter of population biology. Note that this model is commonly known as the logistic model in population biology, but nonetheless specifies a linear decline in the speciation rate. In the paleontological literature, several studies have addressed density-dependent clade growth using logistic models (e.g., Walker & Valentine 1984). However, if the rate of decline in the speciation rate itself declines as the number of species rises, then the exponential model will fit the data better.

We have previously used theory based on the birth-death process (Nee et al. 1994) to fit these models to phylogenetic data (Rabosky 2006a; Rawlings et al. 2008). As a null hypothesis for density dependent diversification, we considered a simple model where the speciation rate varies continuously through time but is independent

of the number of lineages in existence at any point in time. We modelled speciation as a linear function of time:

$$\lambda(t) = \lambda_0\left(1 - t/K\right) \quad (7.3)$$

In this model, which we refer to as the continuous-decline model, the rate of change of the speciation rate is independent of time (e.g., there is no second derivative), and the magnitude of the rate decline increases as $t \rightarrow K$. Although many approaches could be used to model time-dependent speciation rates, we chose this simple model because we felt that it provided a reasonable approximation of monotonic changes in environmental variables through time that might influence diversification rates.

We do not treat diversification under the birth-death process with nonzero extinction for several reasons. First, we have previously shown that patterns of early, rapid diversification as inferred from molecular phylogenies of extant taxa can only be explained by declining speciation rates through time and not by increasing extinction rates (Rabosky & Lovette 2008). Second, high but constant extinction rates will erase the signature of such 'explosive-early' diversification from molecular phylogenies, rendering it impossible to observe even dramatic declines in speciation rates through time (Rabosky & Lovette 2008). For these reasons, extinction rates estimated from phylogenies that appear to undergo temporal declines in diversification rarely differ from zero (Weir 2006; Rabosky & Lovette 2008).

To find the likelihood of phylogenetic data under the continuous-decline model, we used the general probability model developed by Nee et al. (1994); this approach was used to model time-varying speciation and extinction rates in Rabosky & Lovette (2008). Consider a general birth process, where existing lineages give birth to new lineages at a per-lineage, time-varying speciation rate $\lambda(t)$. Let t_i represent the

birth time of each of the N lineages in the phylogeny which survive to the present (time T). The likelihood of the phylogenetic data is given by

$$L = (N-1)!\prod_{i=3}^{N}\{\lambda(t_i)\}\prod_{i=3}^{N}\{\xi_i\}\{\xi_2^2\} \quad , \tag{7.4}$$

where

$$\xi_i = \exp\left[\int_{t_i}^{T} -\lambda(s)ds\right] \quad , \tag{7.5}$$

with t_2 corresponding to the time of the initial bifurcation in the tree. Equation 7.4 is identical to Nee et al. (1994; eqn 20) with no extinction term. This expression considers only $N - 2$ speciation events, because the first two speciation events must have occurred; if they had not, no phylogenetic tree would exist to be observed (Nee et al. 1994). The ξ_2 terms in eqn 7.4 corresponds to these basal branches, and there are two of them.

Equations for modelling time-varying speciation rates were obtained for the continuous-decline model by deriving the appropriate analytical expression for eqn 7.3 in conjunction with eqns 7.4-7.5. Models were fitted to phylogenetic data using Nelder-Mead and BFGS algorithms as implemented in the 'optim' routine for the R programming language. All optimizations were repeated 20 times with random starting parameters to decrease the possibility that solutions reflect local maxima.

Diversification analyses

If ecological opportunity or niche availability facilitated speciation during the radiation of *Dendroica* warblers, we predicted that (i) diversification rates would decline significantly through time, and (ii) models specifying density-dependence of

speciation rates would fit the observed data better than a model where rates decline continuously through time.

We first tested whether previous conclusions about declining diversification rates in the group (Lovette & Bermingham 1999) are robust to the additional data, taxon sampling, and analytical methodologies presented in this paper. We computed the γ statistic (Pybus & Harvey 2000) for ultrametric trees recovered with the BEAST analysis, where $\gamma < 0$ implies decelerating diversification through time. It is well known that incomplete taxon sampling can result in a perceived temporal decline in diversification rates (Nee et al. 1994). Although we included all nominate members of the continental *Dendroica* radiation in our analysis, it is possible that undescribed or morphologically cryptic species could have resulted in a spurious decline in diversification rates over time. To explore the effects of missing species on our analysis, we determined the number of missing lineages that would render the observed γ-statistic insignificant. We assumed that our sample of $n = 25$ lineages represented a proportion f of the true number of lineages and simulated sets of 5000 phylogenies under a pure birth model of cladogenesis for values of f from 0.25 to 1.0.We calculated the γ-statistic for all simulated trees and determined the 0.05 percentile of the distribution of γ for each f; this value corresponds to the lower bound of the 95% confidence interval around the null hypothesis that γ is not significantly less than zero.

We then tested whether the tempo of lineage accumulation during the *Dendroica* radiation is best approximated by density-dependent or continuous-decline models of diversification. We compared the likelihood of the warbler phylogeny under these competing classes of models using the Akaike Information Criterion (AIC). We had no *a priori* predictions as to whether density-dependent diversification should

follow an exponential or linear model. We therefore computed the following test statistic:

(7.6) $$\Delta AIC_{TS} = AIC_{H0} - AIC_{H1}$$

where AIC_{H0} is the AIC score of the null hypothesis model (continuous-decline) and AIC_{H1} is the AIC score corresponding to the best fit hypothesis model (density-dependent exponential or linear). Thus, a positive ΔAIC_{TS} implies that density-dependent models fit the data better than the continuous-decline model. We computed ΔAIC_{TS} for the maximum clade credibility (MCC) tree, which is an estimate of the tree with the maximum *a posteriori* probability. The MCC tree is the tree for which the product of posterior probabilities across all nodes present in the tree is greater than for any other trees in the posterior distribution. To avoid conditioning our results on any particular topology and branch lengths, we computed the distribution of ΔAIC_{TS} over the posterior distribution of trees sampled using MCMC, with the prediction that the continuous-decline model would consistently provide a poorer fit to the data than the density-dependent models.

The analyses described above are critically dependent on the assumption that density-dependent models will not overfit the data in the absence of density-dependent diversification. To test this assumption, we investigated Type I error rates for constant rate phylogenies simulated under both pure birth and continuous-decline models of diversification, assuming both complete and incomplete taxon sampling. For the pure birth model, we simulated 1000 trees of N = 25 taxa under a constant speciation process and tabulated the distribution of ΔAIC_{TS}. To further control for the possibility that incomplete taxon sampling could result in high Type I error rates, we tabulated the distribution of the test statistic for constant-rate phylogenies simulated with

different levels of incomplete sampling (*f*), as described above for the γ-statistic analyses.

We used the method described in Rabosky & Lovette (2008) to simulate phylogenetic trees under the continuous-decline model of diversification (Appendix II). This approach enables continuous-time simulation of phylogenetic trees with time-varying rate parameters. We simulated clade growth under the continuous-decline model assuming both 5-fold and 10-fold reductions in the speciation rate through time. We found parameters of the continuous-decline model (λ_0, K) which would result in an expected value of N = 25 lineages after t = 1.0 time steps, where the expected number of lineages is calculated as

$$n(t) = 2\exp\left(\int_0^{t=1} \lambda(s)ds\right) \tag{7.7}$$

after Nee et al. (1994). We further required that model parameters satisfy the relationships $\lambda_0 = 5\lambda_1$ and $\lambda_0 = 10\lambda_1$ for 5-fold and 10-fold declines, respectively. To simulate incomplete sampling under the continuous-decline model, we found parameters corresponding to 5-fold and 10-fold declines which were expected to result in 33 (f = 0.75), 50 (f = 0.5), or 100 (f = 0.25) lineages at the end of the simulation. Simulated trees were then randomly pruned to the desired sampling level. All phylogenetic simulation was conducted using a modified version of the birth-death tree simulation algorithm from the Geiger package for R (Harmon et al. 2007).

Results

Phylogenetic trees generated under a relaxed clock model of sequence evolution (Figure 7.1) strongly supported previous findings that diversification rates in North American wood-warblers have declined through time. A lineage-through-time (LTT) plot clearly indicates an excess of lineages early in the history of the wood-warbler radiation (Figure 7.2a) relative to the expected rate of lineage accumulation under a constant-rate model of diversification. Calculated γ-statistics for the MCC tree (-3.63) and for the posterior distribution of topologies and branch lengths (2.5% and 97.5% quantiles of -3.63 and -3.19, respectively) indicate highly significant temporal declines in diversification rates ($p < 0.001$). This result is robust to assumptions about missing taxa: γ for the MCC tree is significant even when we assume that our tree contains only 25% of North American wood-warbler species ($p = 0.019$). There is little overlap between the distribution of γ calculated from the posterior distribution of phylogenetic trees and the corresponding null distributions assuming complete and incomplete sampling (Figure 7.2b).

Model-based analyses of diversification provided strong support for density-dependent diversification in wood warblers (Table 7.1). Among the candidate models, the density-dependent exponential model provided the best approximation to the observed pattern of lineage accumulation through time ($\Delta AIC_{TS} = 10.27$). Diversification rates reconstructed using maximum likelihood parameter estimates for this best-fit model ($\lambda_0 = 71$, $x = 1.47$) suggest an explosive burst of diversification early in the history of the radiation (Figure 7.3), followed by a rapid decline in per-lineage diversification rates. In the Appendix II (Figures S2-S3), we provide expected lineage-through-time curves for the three fitted rate-variable models and discuss how differences in these patterns relate to our ability to discriminate between density-dependent and continuous-decline models.

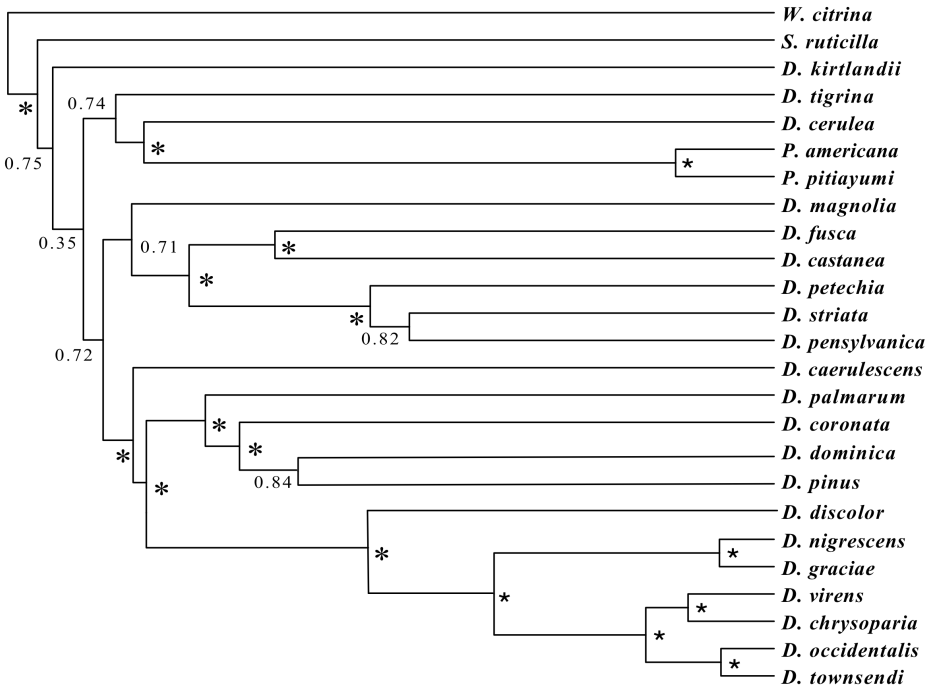

Figure 7.1. Maximum clade credibility (MCC) tree from Bayesian analysis of all continental North American wood-warbler species. Nodes marked with asterisks are supported by posterior probabilities > 0.95. Tree is based on > 9kb of mtDNA and nuclear intron sequence.. Branch lengths are proportional to absolute time.

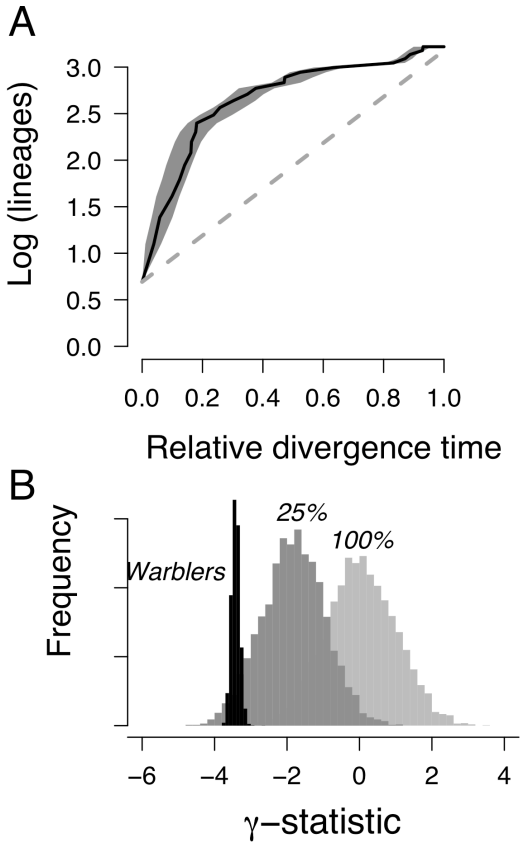

Figure 7.2. (a) Log-lineage through time (LTT) plot for North American wood-warblers. Black line indicates LTT curve for the MCC tree (figure 7.1), and grey shading indicates 95% quantiles on the number of lineages at any point in time as inferred from the posterior distribution of phylogenetic trees sampled with MCMC. Dashed line indicates expected rate of lineage accumulation under constant rate diversification with no extinction. Lineages accumulate quickly in the early phases of the radiation relative to the constant rate diversification model. (b) Posterior distribution of the γ-statistic for wood-warblers (black) in comparison with the corresponding null distributions assuming either complete (f=1) or incomplete (f = 0.25) sampling. Negative values of γ relative to the null distribution indicates decelerating diversification through time.

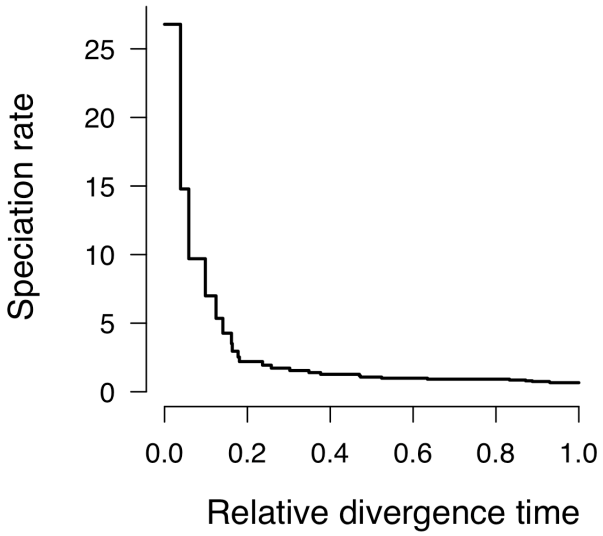

Figure 7.3. Maximum likelihood reconstruction of speciation-through-time curve under overall best-fit model (density-dependent exponential). Rates are given in lineages per-time unit assuming the basal divergence occurred 1.0 time units before the present.

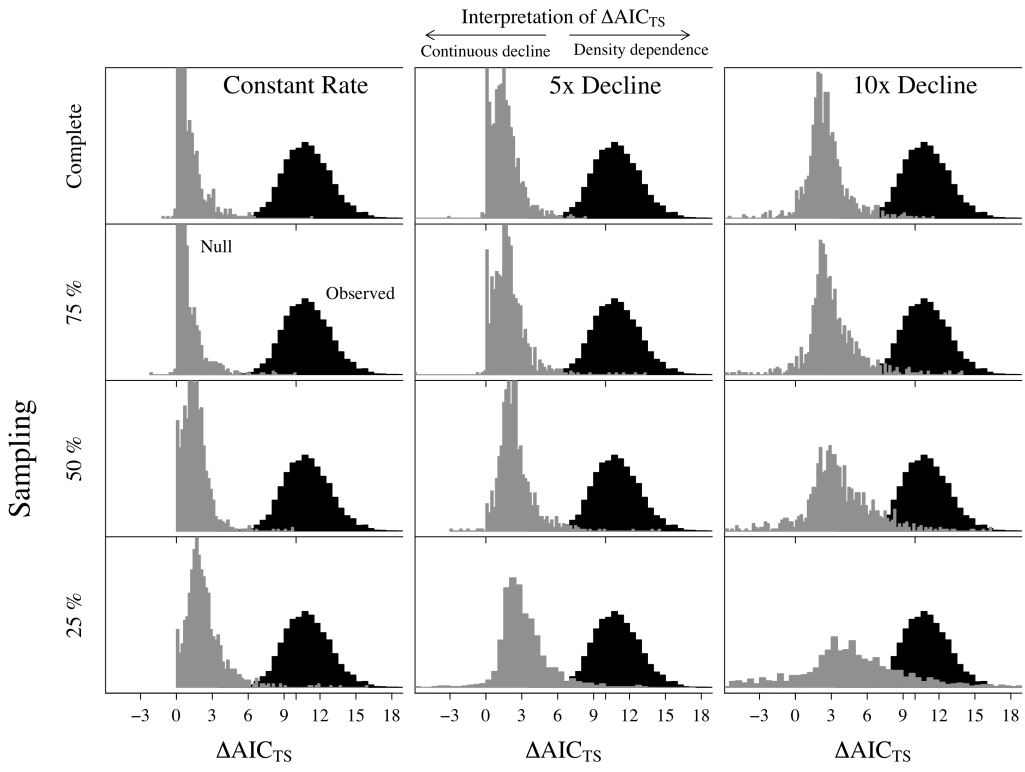

Figure 7.4. Distribution of ΔAIC_{TS} test statistic as tabulated from posterior distribution of wood-warbler phylogenies sampled using MCMC (black). Larger ΔAIC_{TS} values indicate better fit of density-dependent diversification models relative to continuous-decline models. Null distributions of ΔAIC_{TS} statistic (gray) were tabulated from phylogenies simulated under constant rate and continuous-decline diversification models. Null distributions in left column of graphs correspond to constant rate phylogenies; distributions in middle and right columns are phylogenies simulated under 5-fold and 10-fold declines in speciation under the continuous-decline model (eqn 2.3). We further tabulated null distributions under each model assuming taxon sampling was complete or incomplete at 75%, 50%, and 25% levels. Density-dependent models consistently fit the data better than the continuous-decline model, even when large numbers of missing taxa are assumed. The ΔAIC_{TS} statistic for the MCC tree is significant ($p < 0.05$) under all diversification/sampling scenarios except the 10-fold decline with 25% sampling ($p = 0.085$).

Table 7.1. Summary of diversification models fitted to the MCC tree (figure 1) for the North American wood-warbler radiation.

Model	*Log-likelihood*	*ΔAIC*[1]
Density-dependent, exponential	50.47	0
Density-dependent, linear	48.42	4.09
Linear	45.33	10.27
Pure birth	40.45	18.03

[1] Difference in AIC scores between each model and the overall best-fit model.

To test whether our results were robust to uncertainty in phylogeny estimation and assumptions about completeness of taxon sampling, we tabulated the distribution of ΔAIC_{TS} from the posterior distribution of trees and branch lengths sampled using MCMC. We compared this distribution to the null distribution generated by simulating phylogenies under both constant rate and continuous-decline models of diversification with and without incomplete sampling. If ΔAIC_{TS} is characterized by an exceptionally high Type I error rate, we expect density-dependent models to consistently fit the data better than continuous decline models, even when diversification rates have not declined in density dependent fashion.

Under both constant rate and continuous-decline simulations (Figure 7.4), we find that the null distribution of ΔAIC_{TS} is consistently greater than zero. This implies that – on average – density dependent models fit the data slightly better than continuous-decline models, even when data are simulated under a continuous-decline diversification process. However, it is extremely unlikely that our results can be explained by this weak bias in favour of density-dependent models. Although a constant diversification process in conjunction with incomplete taxon sampling can generate the appearance of temporally declining diversification (e.g., Figure 7.2b; null distributions of γ), our results reject the possibility that this artefact underlies the much greater fit of density-dependent models to the wood-warbler data (Figure 7.4, left column). The observed ΔAIC_{TS} statistic for the MCC tree (10.27) indicates significantly greater fit of density-dependent models relative to the null distribution when sampling is assumed to be complete ($p < 0.001$) or incomplete at 75% ($p < 0.001$), 50% ($p < 0.001$), and 25% ($p = 0.008$) levels. Moreover, we find little evidence that a continuous-decline process in conjunction with incomplete sampling (Figure 7.4, middle and right columns) can explain the observed distribution of ΔAIC_{TS} for the wood-warblers. The ΔAIC_{TS} test statistic for the MCC tree is

significantly greater than expected under the null distribution assuming a 5-fold decline in diversification across all levels of incomplete sampling ($p \leq 0.01$). This result is not significant only when we assume at least 10-fold declines in diversification with 25% sampling ($p = 0.085$).

Discussion

We present a new conceptual framework for detecting density-dependent diversification rates and report evidence for this process during the radiation of North American wood-warblers. Although previous studies have suggested that density dependence might account for the apparent deceleration in the rate of cladogenesis observed in molecular phylogenies (Nee et al. 1992; Weir 2006; Phillimore & Price 2008), ours is the first to explicitly test whether patterns of speciation are more consistent with density-dependence than with other processes that can result in temporally declining diversification.

Despite the size of our DNA sequence dataset (9557 combined nucleotides), we were unable to resolve phylogenetic relationships among some early diverging lineages (Figure 7.1; nodes < 0.95 posterior probability). However, our results are robust to this uncertainty in phylogeny estimation. The narrow confidence limits on the reconstructed number of lineages as a function of time (Figure 7.2a) indicates that even alternative phylogenetic relationships among wood-warbler taxa show a similar pattern of lineage accumulation. Likewise, the posterior distribution of γ suggests that virtually all topologies and branch lengths sampled using MCMC show this pattern of early and rapid diversification (Figure 7.2b). The significantly better fit of density-dependent models to the MCC tree (Figure 7.1; Table 7.1) is robust to uncertainty in topology and branch length estimation as well as assumptions about sampling completeness (Figure 7.4). Although incomplete taxon sampling generates a pattern of

lineage accumulation through time that can mimic temporally declining diversification, we find that assuming large numbers of missing taxa does not change our result (Figure 7.4). The ΔAIC_{TS} test statistic shows a bias favouring density-dependent over the continuous-decline model (Figure 7.4); however, more than 75 taxa would need to be missing from our analysis for this to pose a problem under all but the steepest declines in diversification. We consider this degree of incompleteness a highly unlikely scenario given that the North American avifauna is well characterized and many *Dendroica* species have been the focus of densely sampled phylogeographic studies.

Ecological causes of density-dependent speciation

Darwin (1859) commented on the fact that islands often harbour fewer competing species than mainland regions, implying that islands might have "less extermination" of species due to an absence of competition. Such a relaxation of competition in conjunction with an abundance of resources has since been characterized as ecological opportunity (Schluter 2000) and is widely hypothesized to drive both lineage and phenotype diversification during evolutionary radiations (Simpson 1953; Erwin et al. 1987; Foote 1996). This is an intuitively appealing and theoretically plausible model (Walker & Valentine 1984; Gavrilets & Vose 2005) that can explain our finding of density dependent diversification rates in wood-warblers. On a mechanistic level, ecological opportunity can facilitate higher per-lineage rates of speciation by increasing the likelihood that a population will split and successfully occupy multiple peaks on the adaptive landscape.

One alternative to density-dependent speciation is that temporal declines in diversification are attributable to increasing extinction rates during evolutionary radiations. Because the net rate of species diversification through time is the difference

between the speciation rate and the extinction rate, a density-dependent increase in the extinction rate during evolutionary radiations would also generate density-dependent declines in the net diversification rate. There are theoretical reasons why rates of extinction might increase during evolutionary radiations (e.g., Ricklefs & Cox 1972). For example, a limit on total resource availability implies that increasing species diversity will result in lower mean population sizes per species (Levinton 1979; Hubbell 2000). Because population size is a determinant of extinction probability, extinction rates might increase with the number of species. However, time-varying speciation and extinction result in different patterns of lineage accumulation through time (Nee 2001; Weir 2006) and we have previously shown that temporally-declining speciation is the only process that can leave a signature of early, rapid diversification in a molecular phylogeny of extant taxa (Rabosky & Lovette 2008). This result does not rule out the possibility that some evolutionary radiations are characterized by density-dependent extinction rates, but it does suggest that it would be difficult to infer such a process from molecular phylogenies alone.

Ecological opportunity and continental evolutionary radiations

It is widely thought that ecological opportunity might underlie the dramatic ecological diversification of many clades on islands (Baldwin & Sanderson 1998; Lovette et al. 2002) and in freshwater lakes (Bernatchez et al. 1999; Seehausen 2006). It is perhaps unsurprising that colonizing species in these insular environments would encounter a combination of high resource availability and a paucity of competing species, as the formation of these environments results in novel habitats which are characterized – at least initially – by low species richness. However, continental radiations occur against a complex ecological background that differs from comparatively simple island systems, and it is unclear whether the processes and

conditions that facilitate adaptive radiations on islands are also important during continental radiations (Barraclough et al. 1999). It is possible that conditions of ecological opportunity that might exist during the early stages of island radiations generally do not occur in continental systems.

The few quantitative analyses of continental radiations have yielded mixed results: some radiations show patterns of diversification consistent with a role for ecological opportunity (e.g., Harmon et al. 2003; Lovette & Bermingham 1999; Rabosky et al. 2007a; Rabosky et al. 2007b), whereas others do not (McPeek & Brown 2000; Turgeon et al. 2005). Other radiations show patterns of lineage but not phenotype diversification consistent with ecological opportunity, indicating that these two aspects of diversification need not be coupled (Kozak et al. 2006). Still other studies (Irschik et al. 1997) suggest that even closely related taxa can experience different patterns of diversification on continents and islands.

Alternatives to ecological opportunity

Although ecological opportunity has been a favoured explanation for temporal declines in speciation rates (e.g., Weir 2006; Phillimore & Price 2008), these results may also be consistent with other processes that entail no direct relationship between speciation and ecological opportunity. For example, it is possible that species interactions influence various aspects of geographic range size; range size might then influence the probability of allopatric speciation. Theoretical work suggests that species interactions can limit geographic ranges (Case et al. 2005), and there is some evidence that geographic range size is positively correlated with diversification rates (Cardillo et al. 2003). During evolutionary radiations, mean range size might decline as the number of species in a particular biogeographic theatre increases. If declining range size results in lower per-lineage speciation rates, then diversification rates as

inferred from molecular phylogenies could show density-dependence. We note that this model of diversification does not imply that adaptive radiation underlies the temporal declines in speciation commonly observed during evolutionary radiations (e.g., Phillimore & Price 2008), because it allows the possibility that behavioural interference and other interactions unrelated to resource use might drive the pattern.

It is also possible that apparent density-dependence of speciation rates could arise as an artefact of phylogeny reconstruction and branch length estimation. It is well known that underparameterized models of sequence evolution can lead to the impression of temporal declines in diversification (Revell et al. 2005). Although we reconstructed phylogenies using a complex model of sequence evolution with multiple data partitions, it is not clear whether existing models of molecular evolution are sufficient for reconstructing substitutional histories along the deep internal branches of a phylogenetic tree. Rabosky & Lovette (2008) pointed out that 'explosive-early' radiations pose a paradox: evidence from the fossil record suggests that virtually all groups diversify with appreciable background extinction rates, yet high extinction rates render it impossible to observe such rapid radiations in molecular phylogenies. One possible solution is that the pattern is – at least in part – attributable to inadequacy of molecular evolutionary models in general use, and this topic clearly deserves a much more comprehensive treatment.

Summary

We have developed a novel conceptual approach that can distinguish between density-dependent speciation and other processes that result in temporal declines in speciation rates. We do not claim that density-dependence is the only possible explanation for our findings, but our results clearly eliminate two competing alternative scenarios. Patterns of lineage accumulation in North American wood-

warblers are inconsistent with a simple model in which speciation rates vary linearly as a function of time. Moreover, the explosive-early accumulation of lineages is more consistent with density-dependent diversification than an artefactual decline in rates attributable to incomplete taxon sampling.

Previous studies have found co-occurring wood-warbler taxa to differ in both foraging niche and other ecological traits (e.g., MacArthur 1958; Morse 1989; Martin and Martin 2001), and we have previously shown that local warbler assemblages are phylogenetically overdispersed (Lovette & Hochachka 2006). These features suggest the possibility that wood-warbler communities have been assembled through adaptive radiation (Gillespie 2004; but see Freckleton & Harvey 2006). Our finding that speciation rates in wood-warblers show density-dependence adds a novel dimension to our understanding of this continental radiation because it suggests the possibility of coupling between lineage and ecological diversification in this group.

CHAPTER 8

ECOLOGICAL LIMITS ON CLADE DIVERSIFICATION IN HIGHER TAXA

Abstract

Species richness varies dramatically among groups of organisms, yet the causes of this variation remain poorly understood. Variation in species-level diversification rates may partially explain differential species richness among clades, but older clades should also be more diverse, because they will have had more time to accumulate species. Surprisingly, studies that have investigated this question have reached dramatically different conclusions: several claim to find no such age-diversity relationship, whereas a recent and more inclusive study reported that clade age and not diversification rate explains the variation in species richness among animal taxa. Here I address the relationship between clade age and species richness using a model-based approach that controls for variation in clade age and among-clade variation in diversification rates.

Under this approach, the assertion that clade age explains most of the variation in species richness among higher taxa is rejected: in four of five datasets, species richness is effectively independent of clade age. Extreme among-clade variation in diversification rates cannot account for the absence of a positive age-diversity relationship in angiosperms, birds, and teleost fishes. I consider two alternative explanations for these results and find that a clade volatility model positing correlated speciation-extinction dynamics does not underlie these patterns. Rather, ecological limits on clade growth, such as geographic area, appear to mediate temporal declines in diversification within higher taxa.

Introduction

Why do some groups of organisms have so many species and why do other groups have so few? Despite decades of interest in this problem (Raup et al. 1973; Slowinski and Guyer 1989; Sanderson and Donoghue 1996; Mooers and Heard 1997), there is no consensus on the general processes that underlie this pervasive feature of biological diversity. One possible explanation for differential species richness among clades is that it reflects lineage-specific differences in rates of speciation and extinction. These parameters –collectively referred to as diversification rates - are widely known to vary both over time (Sepkoski 1998; Phillimore and Price 2008) and among lineages (Sims and McConway 2003; Coyne and Orr 2004; Davies et al. 2004; Moore et al. 2004) and clearly influence species richness among clades (e.g., Phillimore et al. 2006; Rabosky et al. 2007; Moore and Donoghue 2007). However, another major explanation for these differences in species richness is clade age: older clades have had more time to accumulate species and should be more diverse than younger clades (Labandeira and Sepkoski 1993; Wiens and Donoghue 2004; McPeek and Brown 2007).

For clades that grow with an identical net rate of diversification, there is a strong relationship between extant diversity and clade age (Figure 8.1a), even when speciation and extinction rates are exactly equal (Figure 8.1b). It is therefore surprising that several recent studies have reported nonexistent or even negative relationships between clade age and extant species richness in higher taxa (Magallon and Sanderson 2001; Ricklefs 2006; Ricklefs 2007; Ricklefs et al. 2007). For example, Ricklefs (2006) found a weak but negative relationship between clade age and species richness for avian tribes. This was interpreted as a tendency for clades to grow rapidly to an equilibrium diversity, with comparatively little net diversification occurring after that time. Magallon and Sanderson (2001) found a similar weak effect of clade age on

species diversity in angiosperm higher taxa. In an analysis of major clades of squamate reptiles, Ricklefs et al. (2007) found no effect of clade age on species diversity: older lineages were no more species-rich than young clades.

In contrast to the above-mentioned studies, McPeek and Brown (2007) reported that clade age, but not diversification rate, explains the variation in species richness among animal taxa. They arrived at this conclusion by applying a regression framework to two datasets. First, they considered a set of 163 species-level molecular phylogenies tabulated from the literature, where all phylogenies were required to be time-calibrated and at least 50% complete at the species level. Their second dataset consisted of extant species diversities for a set of animal higher taxa (e.g., teleost and insect orders) with fossil-based estimates of crown clade ages. Because their molecular phylogenetic dataset is composed of comparatively small species-level phylogenies, it addressed a much different temporal scale than the previously mentioned studies that found no relationship between clade age and species richness among plant and animal taxa. They nonetheless reported a significant correlation between age and diversity in their fossil-based dataset of higher taxa ($\rho = 0.65$; $p < 0.001$).

Correlations between age and diversity reconsidered

I obtained the angiosperm crown clade dataset from Magallon and Sanderson (2001) as well as the avian tribes dataset from Ricklefs (2003, 2006). The Spearman correlation between log(diversity) and crown clade age in Magallon and Sanderson's (2001) analysis of major angiosperm clades was negative but nonsignificant ($\rho = -0.12$; $p = 0.77$). Likewise, Ricklefs (2006) found a negative relationship between stem clade age and species richness in avian tribes (Spearman's $\rho = -0.19$; $p = 0.08$). McPeek and Brown (2007) reported a significant relationship between stem clade age

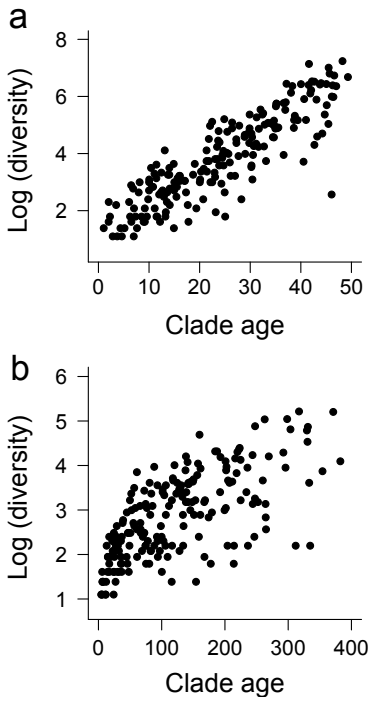

Figure 8.1. Species richness increases with clade age. This is true when the net rate of diversification is positive (*a*; Spearman's $\rho = 0.87$) and even when speciation and extinction rates are equal (*b*; $\rho = 0.69$). Speciation/extinction rates were set to 0.3 and 0.2 respectively for *a*, and 0.1 and 0.1 for *b*. Results for each scenario are based on 200 phylogenies simulated in the R package GEIGER (Harmon et al. 2008); each simulation was conducted for a length of time drawn from a uniform distribution with endpoints of (10, 50) or (25, 400) for *a* and *b*, respectively. Because extinction was present in the simulations, the duration of the simulation does not necessarily correspond to the recovered clade age, as clade age is the time of the basal divergence among the set of taxa that survived to the end of the simulation (e.g., crown clade age).

and species richness across a phylogenetically disparate sample of clades, including orders of teleost fishes, insects, and tetrapods. However, there are significant age differences among the groups they considered ($F_{5, 108} = 29.25$; $p < 0.001$): insects are much older (mean = 149.5; sd = 63.9) than vertebrate orders (mean = 59.6; sd = 39.1). I therefore split the data from this study into constituent groups, assuming that any relationship between age and diversity should also hold among clades within major taxonomic categories.

Because the avian order dataset (McPeek and Brown 2007) overlaps with the much larger avian tribe dataset studied by Ricklefs (2006), I omitted the orders from further analysis. I further eliminated both the amphibian and non-avian reptile datasets, due to low sample sizes (amphibians: n= 3; non-avian reptile: n=4), noting that Ricklefs et al. (2007) found a weak and non-significant relationship between diversity and age in squamate reptiles, using a much larger number of clades (Pearson's $\rho = 0.17$; $p = 0.32$; $n = 36$). Numbers of clades, data source, and Spearman correlations between log(diversity) and age for all groups are shown in Table 8.1.

Only insects show a significant positive relationship between clade age and diversity. In all analyses below, I excluded groups with only one species from datasets with stem clade ages (eliminating 2 of 90 avian tribes) and two species from sets with crown clade ages (eliminating 2 of 20 mammalian orders and 1 of 49 angiosperm clades). These groups were removed because several methods described below require that all clades or lineages have nonzero diversification rates (e.g., crown groups with only two extant species have zero net diversification). Regardless, alternative approaches that included this small number of lineages did not alter the general results presented below.

Table 8.1. Datasets considered in this study, including the number of clades included in the analysis, the Spearman correlation coefficient ρ between (log) species richness and clade age, and data source.

Lineage	*Clades*	ρ	*p*	*Data source*
Angiosperm clades	48	-0.12	0.46	Magallon and Sanderson 2001
Insect orders	25	0.54	<0.01	McPeek and Brown 2007
Avian tribes	86	-0.19	0.08	Ricklefs 2006
Mammalian orders	18	0.04	0.87	McPeek and Brown 2007
Teleost fish orders	29	-0.07	0.70	McPeek and Brown 2007

Possible explanations for weak age-diversity relationships

Several confounding factors that have not previously been considered might weaken the age-diversity relationship in the analysis of higher taxa. For example, the variation in species richness explained by clade age depends fundamentally on the amount of variation in clade age. If there is comparatively little variation in clade age within a sample of higher taxa, then it should be unsurprising if other factors (e.g., diversification rate) explain more of the residual variation in diversity. Likewise, diversification rates clearly vary among clades (e.g., Owens et al. 1999; Chan and Moore 2002; Paradis 2005), and increased variation in rates among clades should weaken the strength of the age-diversity correlation, but there have been no investigations of how rate variation might influence the appearance of 'weak' or 'strong' relationships between age and diversity. Most importantly, both the variance in clade age and the extent of among-clade variation in diversification rates are likely to differ among sets of higher taxa (e.g., avian tribes; angiosperms; teleost orders), making it difficult to generalize and compare the age-diversity relationship across taxonomic groups and phylogenetic scales.

Another possibility entails correlated speciation-extinction dynamics (Gilinsky 1994; Lee and Doughty 2003), whereby clades with the highest net diversification rates tend to have high relative extinction rates. The most rapidly diversifying clades would thus have a high probability of extinction, and any set of clades observed in the present would necessarily reflect this bias. There is considerable evidence from the paleontological literature for correlated speciation-extinction dynamics: high speciation rates are often coupled with high extinction rates (Stanley 1979; Gilinsky 1994; Stanley 2007; Liow et al. 2008). This high rate of both extinction and speciation leads to high rates of lineage turnover through time and results in clade "volatility" (Gilinsky 1994). Under the volatility model, clades with the highest rates of

diversification would be those with the highest relative extinction rates. Clades with high diversification rates should thus be young, because old clades with high diversification rates will likely have gone extinct due to their correspondingly high relative extinction rates. Sepkoski (1998) suggested that such a correlation between speciation and extinction could explain the long-term declines in diversification rates observed in many taxa in the fossil record as an epiphenomenon of the differential extinction of low and high-rate clades.

A final explanation proposes that ecological factors impose constraints on clade growth: clades undergo initially rapid diversification, but these high rates slow through time due to ecological limitations (Valentine and Moores 1972; Rosenzweig 1975). This might reflect density-dependent diversification, where speciation and/or extinction rates are a function of resource availability (Sepkoski 1978; Walker and Valentine 1984; Nee et al. 1992; Price 2008). As the number of species within a particular ecological or biogeographic theater increases, opportunities for speciation decrease or extinction rates increase. This process has long been hypothesized to underlie diversity dynamics in the fossil record (Valentine 1969; Stanley 1973; Rosenzweig 1975; Sepkoski 1978; Carr and Kitchell 1980). Indeed, many researchers from this period (see Rosenzweig 1975; Sepkoski 1978) felt that such diversity dynamics were a logical extension of MacArthur and Wilson's (1963) equilibrium theory of island biogeography to evolutionary processes of speciation and extinction (MacArthur 1969).

I used a model-based approach to test these hypotheses for the absence of a positive age-diversity correlation in higher taxa. I first evaluated the observed correlations between clade age and species richness for all groups in Table 8.1 with respect to a set of null models that make different assumptions about the distribution of rate variation among clades. I then conducted a more detailed analysis of one

dataset, the avian tribes, to test whether the absence of a positive age-diversity correlation can be explained by clade volatility, density-dependent diversification, or diversification mediated by other ecological factors. I used the avian tribes dataset as a model because of the availability of ecological data and the fact that this dataset has been the focus of several previous studies (Ricklefs 2003, 2006).

Models, estimation, and simulation

My general approach to evaluating the raw age-diversity correlations (Table 8.1) was to simulate clades of identical age as the observed clades under variants of a stochastic birth-death process. For each dataset, I generated a null distribution of clade sizes, assuming constant or variable diversification rates among clades. I then computed the Pearson correlation between log(diversity) and clade age. Because this approach maintains the age structure of clades within each dataset, any effects of variation in clade age on the resulting age-diversity correlation is accounted for. This approach also provides an expectation for the correlation between age and diversity under alternative models of diversification rate variation among clades. I conducted 5000 simulations per dataset and null model. Clade ages used in this study were derived from DNA hybridization data (avian tribes) and the stratigraphic record (insects, fish, mammals, angiosperms) and uncertainty in the estimates is unknown. The analyses below assume that error in clade age alone does not account for the poor relationship between age and extant species richness.

I used models based on a general birth-death process, where each lineage gives rise to new lineages with per-capita rate λ and goes extinct with rate μ (Kendall 1948; Nee et al. 1994; Rabosky and Lovette 2008b). These parameters in turn specify the net diversification rate r, and the extinction fraction ε, where $r = \lambda - \mu$ and $\varepsilon = \mu/\lambda$. These latter parameters are especially important and specify the distribution of speciation

times in a reconstructed phylogenetic tree, the expected number of descendants of a diversification process, as well as the long-term probability of lineage extinction (Nee et al. 1994; Rabosky 2006) – at least for the time-homogeneous diversification process.

I considered three alternative models of variation in diversification rates among clades. In the first, I assumed that the net diversification rate *r* was constant among lineages. In the second model, I relaxed this rate-constancy by assuming that r for each clade was drawn from a gamma distribution. Finally, I considered an overdispersed rate model, where rates are drawn from a uniform distribution (see below). The relaxed rate model thus assumes a lower variance for the distribution of rates than the uniform distribution, which results in the greatest range and variance in clade size. For each model, I considered three relative extinction scenarios: $\varepsilon = 0$, $\varepsilon = 0.95$, and random ε. In the latter model, I drew ε for each clade uniformly on the interval [0, 1) during each replicate simulation. Thus, under the random model, each clade within each dataset (e.g., angiosperms) would have a potentially unique ε value.

Constant net diversification rate

Under the constant rate model, I simulated clade diversities given an estimate of the net diversification rate for each dataset, holding this parameter constant among clades. The estimator was found by maximizing the likelihood function

$$L = \prod_{i=1}^{N} \Pr\left(n_i | t_i, \varepsilon_i, r\right) \tag{8.1}$$

where N is the total number of clades and n_i, t_i and ε_i are the diversity, age, and relative extinction rate of clade i. Note that ε_i varies only under the random ε model. The probability of k lineages is calculated as

$$\Pr\left(N(t)=k\right)=\sum_{j=0}^{\min(a,k)}\binom{a}{j}\binom{a+k-j-1}{1}\alpha^{a-j}\beta^{k-j}\left(1-\alpha-\beta\right)^{j} \tag{8.2}$$

(Bailey 1964; Raup 1985), where a is the number of ancestral lineages

$$\beta=\frac{\exp\left(rt\right)-1}{\exp\left(rt\right)-\varepsilon} \tag{8.3}$$

and $\alpha = \varepsilon\beta$. For clades with stem ages, $a = 1$, because the clade age represents the time required for a birth-death process beginning with a single lineage (e.g., immediately after the clade ancestor split from its sister group) to diversify into k lineages. For clades with crown group ages, $a = 2$, because the clade age corresponds to the time of the basal bifurcation in the crown group; at this time, exactly two lineages were alive which left descendents to be observed in the present. Finally, we condition equation 2 on the probability that the stem or crown group ancestor(s) survived to the present, as otherwise they would not exist to be observed. In the case of the stem ancestor, this probability is simply $1 - \alpha$, and for crown group ancestors, $1 - \alpha^2$.

Under the random ε variant of this model, the net diversification rate is constant among lineages, but λ and μ are not. This simple equivalence of net diversification rates among lineages with different values of ε does not imply that they will produce the same expected number of lineages per unit time. Indeed, as the relative extinction rate rises, the net diversification rate required to produce a given diversity level drops (Magallon and Sanderson 2001; Appendix III, Figure S1). It is thus difficult to compare 'diversification rates' among clades without assuming homogeneous ε among lineages.

"Relaxed" net diversification rate

The relaxed rate model assumes that clade rates are drawn from an overall distribution of rates. This is conceptually similar to a framework used to relax the assumption of constant substitution rates among lineages in molecular evolutionary studies (e.g., Thorne et al. 1998; Drummond et al. 2006). I first approximated the distribution of diversification rates among clades within datasets using the stem or crown clade estimators of *r* from Magallon and Sanderson (2001; eqns 6-7). Strictly speaking, this is the observed distribution of rates only if we make the assumption that rates for each clade are fully independent, but this is a conservative assumption in the context of this analysis. Note again that for the random ε model, *r* for each clade was estimated assuming a uniquely drawn value of ε, and each replicate simulation entailed drawing new ε parameters for each clade. I assumed rate variation among clades followed a gamma distribution, after comparing log-likelihood values for the data under gamma and other potential candidate distributions, including the lognormal. I found maximum likelihood estimates of scale and shape parameters given the estimated rates for each clade. Simulations were conducted by drawing *r* for each clade from a gamma distribution parameterized to fit the observed data, thus mirroring the observed heterogeneity in rates among clades within datasets. Clades were then simulated given these diversification parameters. As above, the process was repeated assuming $\varepsilon = 0$, $\varepsilon = 0.95$, and random ε.

Overdispersed net diversification rate

For the overdispersed rate model, net diversification rates were drawn from a uniform distribution. Upper and lower bounds were determined by the observed distribution of rates as calculated above; I arbitrarily set the endpoints of the distribution by calculating the tenth percentile of the range of observed rates and

subtracting this value from the observed minimum to obtain a lower bound, or adding the value to the maximum to obtain the upper bound. If the lower bound calculated in this fashion was less than zero, it was reset at half the distance between the observed minimum and zero. This model results in a much greater variance in rates and diversities among clades than both the actual data and relaxed rate model (Appendix III, Figure S2).

Simulations

From the basic probability model for the birth-death process (Kendall 1948; Nee etal. 1994), we note that the probability of k lineages at time t, conditional on survival of the process to the present, is

$$\Pr\big(n(t)=k\big)=\big(1-\beta\big)\beta^{k-1} \quad , \quad k>0 \tag{8.4}$$

This is simply a geometric distribution with parameter $1-\beta$. It follows that drawing from this distribution is a simple method for simulating diversity given r, ε, and t. For groups with a single ancestral lineage, we simply draw from this distribution until a non-zero result is obtained. For groups with two ancestral lineages, we draw two nonzero numbers from the distribution; their sum is a single simulated diversity value. All analyses and simulations were conducted in the R programming environment.

Clade volatility analyses

I conducted simulations to test whether the clade volatility model could generate a negative correlation between species richness and clade diversity. I considered two general models: ε correlated with r, and ε correlated with λ. For each model, I first drew 10^4 stem clade ages with replacement from set of stem group ages

for the avian tribes dataset. I then associated each stem age with a random ε value, with ε drawn uniformly from [0.2, 1.0). For the correlated ε-r model, I drew 10^4 random r values on the interval (0.02, 0.78); these values correspond to the 2.5% and 97.5% quantiles of the estimated distribution of r for the avian data when ε is assumed to be uniformly distributed on [0.2, 1.0). The r values were sorted and ranked and each r was paired uniquely with the ε value with the corresponding rank. Thus, the largest r value was associated with the largest ε, the smallest r with the smallest ε, and so on. There was thus a perfect correlation between ε and r. For the correlated ε-λ model, I repeated this procedure, but drew λ uniformly on (0.07, 6.8). Each clade thus had a start time for diversification and unique ε-r or ε-λ combination. These ranges for r and λ correspond to ranges estimated for the actual avian dataset.

The simulation consisted of three steps. First, I calculated the probability that each clade would survive to the present (time T) as

$$\Pr(k > 0) = \frac{1-\varepsilon}{1-\varepsilon\exp(-rT)} \tag{8.5}$$

after Kendall (1948); extinction was then simulated by drawing a random number uniformly on the interval (0, 1). If this number exceeded the survival probability, the clade went extinct. For the set of surviving clades, I simulated clades using draws from the geometric distribution as described in the *Simulations* subsection. I then tabulated the bootstrap distribution of age-diversity correlation coefficients by sampling sets of 90 clades (the number of clades in the avian dataset) and computing the correlation between log(diversity) and clade age.

Ecological diversification models

To test whether ecological factors can account for the absence of a positive age-diversity correlation, I constructed several models of time-varying diversification with and without ecological covariates and fitted these models to the avian tribe data. Because geographic area is predicted to exert a substantial influence on clade diversity (Sepkoski 1976; Losos and Schluter 2000), I used the geographic region occupied by each clade as a simple index of "ecology"; areas occupied by each tribe were taken from Ricklefs (2006). Geographic area corresponds to the area of the major zoogeographic regions in which extant members of each clade occur. The first model proposes that the net diversification rate r declines exponentially through time, as expected under some models of density-dependent diversification (Nee et al. 1992; Rabosky & Lovette 2008a):

$$r_t = \lambda_0 e^{-zt} - \varepsilon\lambda_0 e^{-zt} \tag{8.6}$$

$$= \lambda_0 e^{-zt}\left(1-\varepsilon\right)$$

Here λ_0 represents the initial speciation rate, z specifies the rate of decline through time, and t is time measured from the start of each radiation (rather than absolute time); thus, each clade begins diversifying with the same initial high rate. Note that extinction (μ) appears in this expression as $\varepsilon\lambda(t)$, where $\lambda(t) = \lambda_0 e^{-zt}$. I extended this model to include an ecological covariate, such that diversification rates were scaled by the log of the geographic area occupied by each clade:

$$r_{i,t} = c\log\left(A_i\right)e^{-zt}\left(1-\varepsilon\right) \tag{8.7}$$

where $r_{i,t}$ is the net diversification rate of clade i at time t, where A_i is the area occupied by clade i. I refer to the models in eqn 8.6 and 8.7 as density-dependent models, even though we are not explicitly modeling declining diversification rates as a function of the number of lineages in existence. This assumes that geographic area has been constant through time, but this is a conservative assumption, because time-varying geographic area should only make it more difficult to detect a true area effect. I also modeled diversification as a constant-rate but area-specific process, such that net diversification rates are proportional to the area occupied by each clade:

$$r_i = c\log(A_i)(1-\varepsilon) \tag{8.8}$$

For comparison, I also considered a model of time-varying diversification, such that the overall diversification rate increases or decreases linearly through time:

$$r = (\lambda_0 + k\tau)(1-\varepsilon) \tag{8.9}$$

where τ is absolute time. For the non-homogeneous birth-death process, we compute β (eqn 8.4) as

$$\beta = 1 - \exp\left[\rho(T,t_0)\right]P(t_0,T) \tag{8.10}$$

where

$$P(t,T) = \left[1 + \int_t^T \mu(\tau)\exp\left[\rho(\tau,t)\right]d\tau\right]^{-1} \tag{8.11}$$

and

$$\rho(T,t) = \int_{t}^{T} \{\mu(\tau) - \lambda(\tau)\} d\tau \quad (8.12)$$

after Bailey (1964; eqns 9.19 – 9.35). I found exact solutions to ρ(t, T) for the models given by eqns 8.6-8.9. Numerical integration of 8.11 was performed using a Gaussian quadrature routine, and parameter estimates were found by maximizing eqn 8.1, where the probability of k lineages is given by eqn 8.4. All optimization was performed using Nelder-Mead and L-BFGS-B methods as implemented in the 'optim' routine for R. I compared the fit of the models using the AIC. I then simulated 5000 sets of clades of the same age as the avian tribes under the maximum likelihood parameter estimates for each model to test whether these models could generate weak or absent correlations between age and diversity. Simulations were conducted by drawing clade sizes from a geometric distribution (eqn 8.4), with parameter 1 - β computed for the inhomogeneous birth-death process (Bailey 1964).

Results

Age-diversity correlations and among-clade rate variation

Clade age is a poor predictor of species richness: with the exception of insects, the observed age-diversity correlation is much lower than expected when clade sizes are simulated using maximum likelihood estimators of the net diversification rate (Figure 8.2, upper row). For teleosts, birds, and angiosperms, clade age explains none of the variation in species richness ($\rho < 0$; Table 8.1), unless we allow the possibility of an inverse relationship between age and diversity. These results are robust across all three extinction scenarios; although Figure 8.2 shows results for $\varepsilon = 0$, both $\varepsilon = 0.95$ and random ε give similar results (Appendix III, Table S1). Numerical results for all

simulations (mean, standard deviation, and other simulation parameters) are given in Appendix III, Table S1.

These general results hold after relaxing the assumption of homogeneous diversification rates among clades (Figure 8.2, middle row). When gamma-distributed heterogeneity in r is incorporated into simulations, the observed age-diversity correlations are still lower than the expected correlations for all datasets. This result is significant across all extinction scenarios for teleosts, birds, and angiosperms. While not significant for insects and mammals, the results trend in the same direction.

Even simulations conducted with overdispersed, uniformly distributed net diversification rates generally fail to recover age-diversity correlations as low as those observed in mammals, teleosts, angiosperms, and birds (Figure 8.2, lower row). Observed correlations are significantly lower than simulated correlations for teleosts, birds, and angiosperms ($p < 0.05$), and trend in this direction for mammals ($p < 0.15$). Only in insects is there any evidence that such rate overdispersion is a possible explanation for the weak correlations between age and diversity. As in constant rate and relaxed rate simulations, these results are robust across the three extinction scenarios considered (Appendix III, Table S1).

Clade volatility analyses

For both the correlated ε-r and ε-λ models, a majority of clades went extinct during the simulation (56.9% and 55.3%, respectively). Moreover, extinction had profound consequences for the distribution of relative extinction ratios among surviving clades (Figure 8.3a), as expected. However, even this considerable and selective clade mortality failed to eliminate the positive age-diversity correlation (Figure 8.3b, 8.3c). Despite a perfect correlation between ε and r, and between ε and λ, age continues to exert a potent effect on species richness.

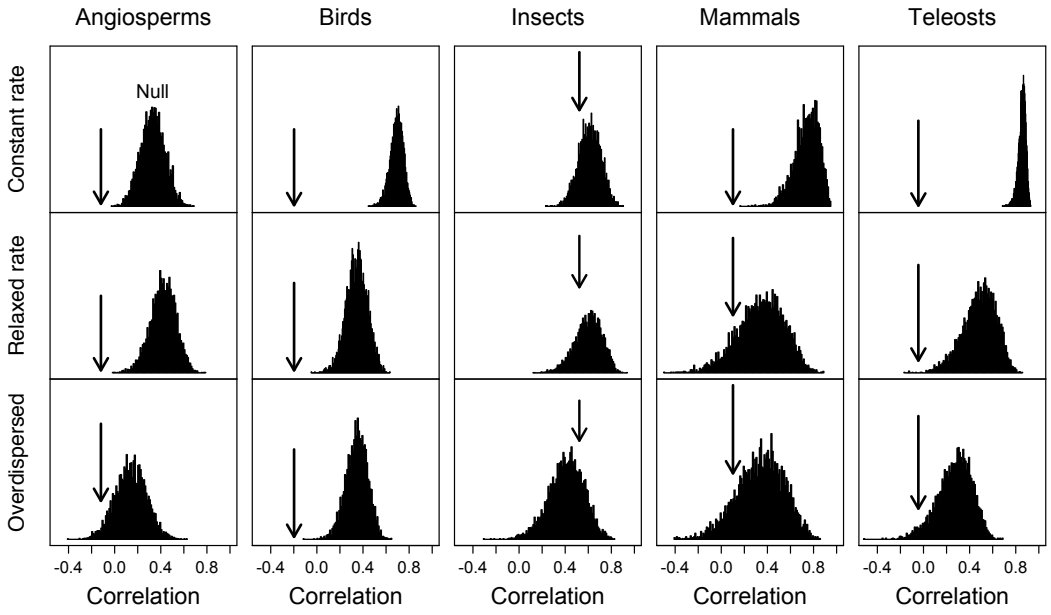

Figure 8.2. Frequency distributions of the correlation between log(diversity) and age for clades simulated under constant rate (top row), relaxed rate (middle), and overdispersed rate (bottom) models. Observed correlations for the five datasets are indicated by arrows. Results shown only for $\varepsilon = 0$ but are very similar to those obtained for $\varepsilon = 0.95$ and random ε models (Supplementary Table S1, Appendix III). All simulations for a given taxonomic group (e.g., Angiosperms) used the observed set of clade ages; the distribution of correlations thus indicates the expected correlation under a stochastic birth-death process. Correlations for angiosperms, birds, mammals, and teleosts are significantly less than expected under the constant rate model; and correlations for angiosperms, teleosts, and birds are significantly less than expected under all three simulation models. Mean correlations, effect sizes, and other simulation parameters are given in the Supplementary Table S1. Results for each group are based on 5000 simulated datasets.

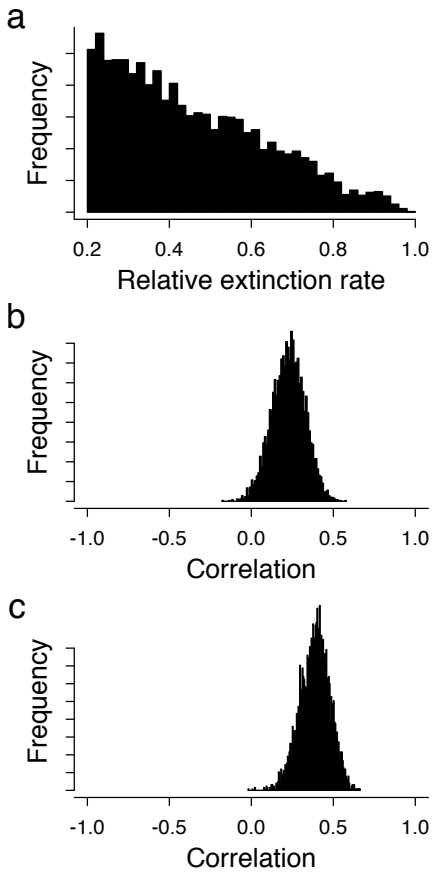

Figure 8.3. Clade volatility simulations indicate that tightly-coupled speciation-extinction rates cannot explain the lack of correlation between clade age and species richness. (a) Distribution of relative extinction rates among surviving clades in the simulation with correlated ε-r dynamics. Distribution was uniform among 10^4 clades that started the simulation, but more than half went extinct, and those that went extinct tended to have high ε values. (b) The correlation between age and species richness for simulations assuming a perfect correlation between ε and r, and (c) between ε and λ. Even this perfect correlation between extinction probability (ε) and net diversification or speciation results in a substantial positive correlation between species richness and clade age.

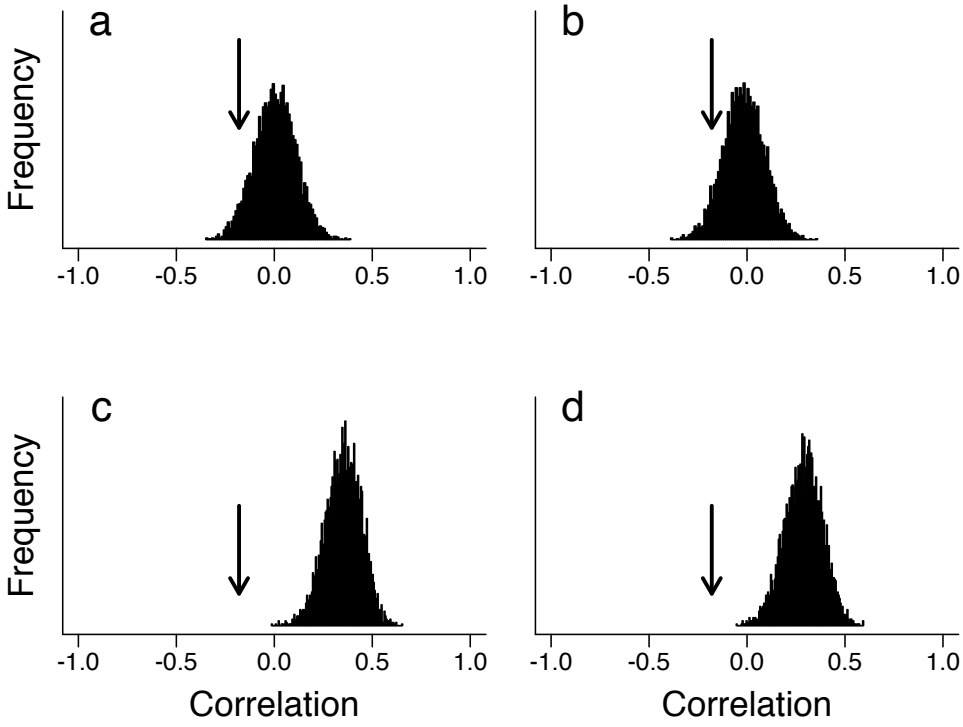

Figure 8.4. Distribution of correlations between log(diversity) and age for clades simulated under (a) density-dependent, (b) density-dependent with geographic area, (c) geographic area only, and (d) linear rate change models. Simulations were parameterized with maximum likelihood parameter estimates for the avian tribes dataset. Arrows indicate the observed age-diversity correlation for the avian data. Both density dependent models (a, b) generate age-diversity correlations consistent with the observed data, whereas the area-only (c) and linear rate change models (d) generate positive age-diversity relationships. Results shown are for simulations assuming $\varepsilon = 0.95$ but are virtually identical to those obtained for other values of ε.

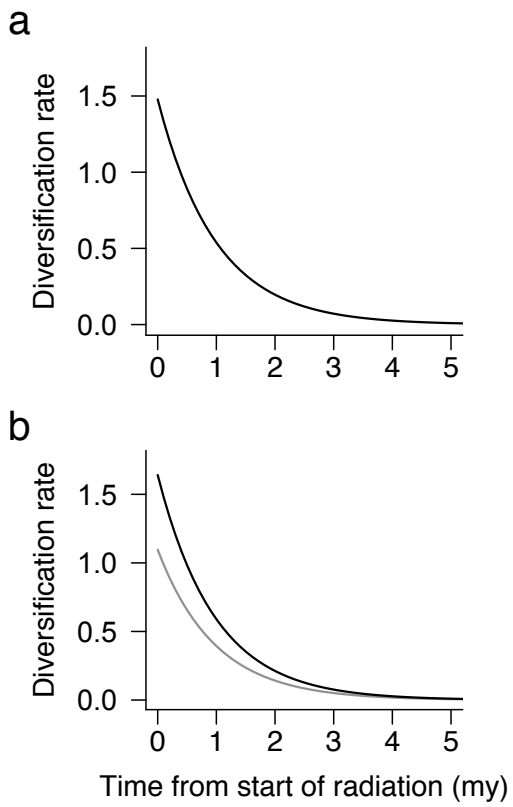

Figure 8.5. Maximum likelihood estimates of net diversification-through-time curves (λ - μ) for the avian tribes dataset under density-dependent models. (a) Simple exponential decline in rates (eqn 8.6). (b) Exponential decline in rates with geographic area occupied by each clade as a covariate (eqn 8.7). Black curve indicates net diversification rate-through-time curve for a clade occupying the largest geographic region among the avian tribes data (cosmopolitan); gray indicates the curve for the clade with the smallest distribution (New Zealand). In both (a) and (b), almost all diversification is concentrated within the initial three million years of the radiation.

Table 8.2. Summary of ecological and non-ecological models fitted to the avian tribes dataset.

Model	n_p [1]	*LogL*	ΔAIC [2]	*Parameters*
$\varepsilon = 0$				
Constant rate	1	-529.1	135.1	$\lambda = 0.57$
Density dependent (eqn 8.6)	2	-468.3	15.5	$\lambda_0 = 4.47$; z = 1.06
Density-dependent + area (eqn 8.7)	2	-460.5	0	c = 0.248; z = 1.01
Area only (eqn 8.8)	1	-526.3	129.6	c = 0.034
Linear (eqn 8.9)	2	-507.1	93.2	$\lambda_0 = \sim 0$; k = 0.04
$\varepsilon = 0.95$				
Constant rate	1	-486.5	50.1	$\lambda = 0.19$
Density dependent	2	-468.3	15.6	$\lambda_0 = 29.55$; z = 1.01
Density-dependent + area	2	-464.1	7.3	c = 1.75; z = 1.02
Area only	1	-474.2	25.5	c = 0.182
Linear	2	-478.6	36.1	$\lambda_0 = \sim 0$; k = 0.27

[1] The number of parameters in each model
[2] Difference in AIC scores between each model and the overall best-fit model (density-dependent + area, $\varepsilon = 0$).

Ecological diversification models

Comparison of model likelihoods in an AIC framework clearly indicates that the density-dependent model with geographic area fits the avian tribe data best (Table 8.2), under both ε = 0 and ε = 0.95. The next best model was a simple density-dependent model without incorporating geographic area (ΔAIC = 15.5, $\varepsilon = 0$; ΔAIC = 15.6, $\varepsilon = 0.95$).

To assess whether any of the models could recover the age-diversity relationship observed for birds, I simulated clades under the maximum likelihood parameter estimates for each model, using the observed set of avian stem clade ages. The density-dependent models result in a breakdown of the age-diversity relationship (Figure 8.4a, b), whereas the area-only and linear change models (Figure 8.4c, d) recover a substantial correlation between log(diversity) and clade age. These results indicate that the two density-dependent models, uniquely among all other scenarios considered, can account for the absence of a correlation between log(diversity) and clade age. The fitted net diversification through time curves for the density-dependent models (Figure 8.5) suggest that increases in clade species richness generally occurred within the first 3 million years (my) from the start of each radiation. Diversity levels appear not to have increased with time beyond this point, although short-term fluctuations would likely have occurred. Because the minimum estimated stem clade age for the avian tribes is 3.5 my (mean = 8.2 my), this suggests that the total species diversification experienced by any clade is effectively independent of the age of the clade.

These results also suggest that estimates of extinction are model-dependent: the constant rate, area-only, and linear change models all provide strong support for high relative extinction rates (Table 8.2). The difference in AIC scores between $\varepsilon = 0$ and $\varepsilon = 0.95$ for these models ranges from 57.1 to 104.1. However, simulations

indicate that these models result in positive age-diversity correlations, even under high relative extinction rates. In contrast, the density-dependent models provide no support for high relative extinction rates and further suggest that power to distinguish between alternative extinction scenarios is low. The density-dependent model with geographic area as a covariate provides modest support for $\varepsilon = 0$, but the variant without area is unable to distinguish between low and high ε values. Yet it is only these latter models which are capable of recovering the observed relationship between clade age and diversity (Figure 8.4).

Discussion

A number of surprising results follow from these analyses. While previous studies painted a conflicting picture of the relationship between age and diversity (Ricklefs 2006; McPeek and Brown 2007), this analysis clearly indicates that clade age alone explains little of the variation in species richness among recognized higher taxa. When clades are simulated under a constant rate birth-death process, the correlation between age and species richness is far greater than that observed in four of five datasets considered here (Figure 8.2; Appendix III, Table S1). Species richness can be thought of as the outcome of four factors: clade age, diversification rate, susceptibility to mass extinction events, and the stochasticity of the diversification process. Because the simulation procedure I used is inherently stochastic and explicitly controls for variation in clade age, these results imply that variation in diversification rates, either over time or among lineages, must account for the failure of the constant rate diversification model to explain the data.

However, the absence of a relationship between age and species richness does not merely result from heterogeneity in diversification rates among clades. I predicted that the magnitude of the correlation between clade age and species richness would

decline as rates varied among clades. Although the strength of the correlation tends to decay with increased rate heterogeneity (Figure 8.2; Appendix III, Table S1), contrary to my expectation both the relaxed and overdispersed rate models were generally insufficient to explain the observed diversity-age correlations. Variation in rates among clades is simply insufficient to eliminate the strong effect of age on species richness.

If the absence of an age-diversity relationship cannot be attributed to heterogeneous diversification rates among clades, I proposed that clade volatility or ecological controls on diversification might underlie the pattern. Under the clade volatility model, extinction rates might be tightly coupled to speciation rates, and clades with high diversification rates would be more likely to go extinct. Thus, among the set of clades that have survived to the present, only the youngest clades would be likely to have a high rate of diversification, because older clades with high diversification rates would likely have gone extinct.

My simulations indicate that such a clade volatility scenario is unlikely to explain the weak/negative age-diversity correlation among higher taxa. The simulation scenario presented here corresponds to 'perfect volatility': higher r or λ values are always associated with higher ε values, and there is no variation in this relationship. Yet even with this idealized relationship, extinction of high-rate clades does not obscure the effects of age on species richness. It is likely that 'true' distributions of ε, r, and λ are much more complex than those employed in the simulation model, and clade volatility might yet be important for some yet-unexplored regions of parameter space. However, incorporating this additional realism is a daunting task: it is difficult enough to quantify real-world variation in these parameters among surviving clades (Nee 2001; Paradis 2004), let alone the set of clades that have not survived to the present.

The most likely explanation for the absence of an age-diversity correlation is that ecological factors impose constraints on clade growth: clades undergo initially rapid diversification, but these high rates slow through time due to ecological limitations which almost certainly include geographic area (Valentine and Moores 1972; Rosenzweig 1975). I found that simple models of temporally declining diversification provided a better fit to the data relative to other models of among-clade and time-varying diversification (Table 8.2). Most importantly, only these scenarios were able to recover a correlation between log(diversity) and clade age consistent with the avian tribes data (Figure 8.4).

These results enable us to discriminate between two ecological models of diversification. Under the first model, net diversification rates vary as a function of ecological factors (e.g., geographic area) but remain relatively constant through time. This might occur if opportunities for speciation arise in proportion to the geographic area occupied by a clade, without necessarily positing hard limits on total diversification. The second model proposes that there is a limit to total diversification, with the limit set by geographic area and presumably, other ecological factors. Although these models are not mutually exclusive, comparisons of model fits (Table 8.2) and simulation results (Figure 8.4) clearly imply that limits on diversification account for the dominant signal in these data. If clade diversity is set by ecological factors and diversification rates decline rapidly (Figure 8.5), then attempts to measure and compare diversification rates among higher taxa may be severely compromised if rates are assumed to have been constant through time within clades.

It is also interesting that the density-dependent diversification scenarios suggest either low relative extinction rates or that power is very low to discriminate between high and low relative extinction rates (Table 8.2). This conflicts directly with the results of previous studies (Ricklefs 2006; Ricklefs et al. 2007), which suggested

high relative extinction rates for higher taxa as inferred from clade age and species richness data. Moreover, this pattern is contrary to what is seen for constant rate, area-specific, and linear models (Table 8.2), all of which strongly favor high relative extinction rates. That none of the latter models can reconstruct an age-diversity relationship consistent in any way with the actual data (Figure 8.4) suggests that all three represent "wrong models", regardless of their ability to fit the data as measured by the AIC. Had we not considered the density-dependent models, we would have concluded that the area-specific model with $\varepsilon = 0.95$ fit the data best, when this model still leads to a positive age-diversity relationship. The difference in fit between these models under $\varepsilon = 0$ and $\varepsilon = 0.95$ is considerable (Table 8.2), suggesting that use of inappropriate models can lead to a remarkably high but illusory level of confidence in estimated extinction rates.

The idea that density-dependent diversification might play a major role within species-level radiations is supported by recent analyses of diversification patterns within molecular phylogenies of extant taxa (Lovette and Bermingham 1999; Pybus and Harvey 2000; Harmon et al. 2003; Ruber and Zardoya 2005; Weir 2006). Two recent meta-analyses found widespread evidence for temporal declines in speciation during species level radiations (Phillimore and Price 2008; McPeek 2008). Rabosky and Lovette (2008a, b) found that patterns of diversification in North American wood-warblers were better explained by density-dependent speciation than by other factors that could cause real or apparent declines in the rate of speciation through time. These results have often been interpreted as a consistent with an adaptive radiation model of diversification, whereby rates of speciation slow through time as ecological niches become progressively filled during the course of a radiation. The results of this study suggest that similar temporal declines in diversification rates may be the dominant signal in higher taxa as well.

Although 'niche filling' models have an obvious relevance to understanding temporal declines in diversification within depauperate insular environments, their role in explaining the pattern in ecologically complex continental settings is far from clear. We have no reason to believe that ecological niche availability was any higher during the early evolution of the higher taxa considered in this study. One possible explanation is that higher taxa are recognized as such precisely because they have acquired phenotypic and ecological traits that enabled them to diversify within a novel 'adaptive zone' (Simpson 1953; Stanley 1979). The acquisition of such traits might set the stage for a rapid burst of diversification, but this diversification would slow through time as ecological space becomes progressively more saturated during the course of the radiation. Under this model, the origins of higher taxa should be associated with expansions of total ecological space. A related model is that higher taxon origination is associated with the evolution of competitively superior phenotypes, such that one clade undergoes rapid diversification at the expense of other clades, with little net change in diversity and no expansion of ecological space.

One possible alternative explanation for the pattern is that the absence of an age-diversity correlation is an artifact of error in the estimation of clade age, species richness, or both. For example, if there is no real variation in clade age among a set of taxa, then error in the estimation of clade age can suggest the absence of a positive age-diversity correlation, when the phenomenon is attributable solely to error in age estimation. However, Ricklefs (2006) conducted simulations which suggested that this was an unlikely explanation for the avian tribes dataset; he reported that error of at least 22% in the estimation of clade ages would still lead to a positive age-diversity relationship.

Another alternative is that evolutionary radiations frequently experience an "overshoot" in species richness, such that young radiations have an excess of species

relative to older radiations (Gavrilets and Vose 2005; Price 2008). Extinction then causes this excessive diversity to relax over time to levels seen in other geographic areas. This could possibly lead to a weak age-diversity relationship, because young clades are expected to be proportionately more speciose than older clades. It is unclear whether this phenomenon may operate over timescales relevant to this study, and the generality of such an overshoot effect likewise remains poorly known.

In conclusion, the results of this study indicate that extant species richness is frequently decoupled from clade age in higher taxa and suggest that the pattern reflects ecological constraints on clade growth. Such temporally declining diversification could lead to a dynamic balance between speciation and extinction rates, resulting in zero net diversification after achieving equilibrium (Ricklefs 2007); the equilibrium point might in turn be determined by a variety of resource-related, geographic, or other regional factors. That such factors might exert a strong control on total diversification is supported by the large number of studies that have demonstrated a positive relationship between geographic area and clade diversity (e.g., Sepkoski 1976; Losos and Schluter 2000; Davies et al. 2005; Ricklefs 2006; Ricklefs et al. 2007).

APPENDIX I

SUPPLEMENT TO CHAPTER 6

Table s1. Major *Ctenotus* species groups, after Storr et al. (1999).

Ctenotus *species group (morphology)*	*Species included*
atlas	*atlas, piankai, quattuordecimlineatus, xenopleura*
australis	*australis, fallens, robustus, saxatilis, spaldingii*
colletti	*calurus, leae, nasutus*
essingtonii	*essingtonii, hilli*
grandis	*grandis, hanloni*
labillardieri	*labillardieri, youngsoni*
leonhardii	*gagudju, hebetior, leonhardii, septenarius, maryani, orientalis*
pantherinus	*angusticeps, pantherinus*
quinkan	*terrareginae*
rubicundus	*rubicundus*
schevilli	*astarte, schevilli*
schomburgkii	*brooksi, schomburgkii, strauchii*
Uncertain	*taeniolatus*

Table s3. Comparison of Bayesian, likelihood, and parsimony bootstrap support values for major clades discussed in text.

	Bayesian	Likelihood[a]	Parsimony[b]
Ctenotus monophyly	100	100	97
Lerista monophyly	100	96	96
Ctenotus + *Lerista* monophyly	100	85	86
Monophyly of Australian sphenomorphine skinks	100	98	87
Monophyly of Australian sphenomorphines excluding *Notoscincus*	100	79	75

[a]Likelihood bootstrap analyses (1000 replicates) were conducted using the distributed phylogenetics platform Multiphyl (Keane et al. 2007; http://distributed.cs.nuim.ie/multiphyl.php). Heuristic tree searches used nearest-neighbor interchange (NNI) with initial trees determined by neighbour-joining. Model selected by Multiphyl was GTR + G + I, with four gamma rate categories.

[b]Parsimony boostrap analyses (1000 replicates) were implemented in Paup* 4.0b10 (Swofford 2002). We used a heuristic search algorithm with tree-bisection-reconnection (TBR) branch swapping and 100 random taxon-addition-sequence replicates per bootstrap replicate. All characters were given uniform weighting.

Table s4. Combined probability of observed standing diversity levels for lineages in the sphenomorphine skink radiation under the null hypothesis that diversification rates for each lineage are equal to those of the radiation considered as a whole. This is a conservative test for among-lineage heterogeneity in diversification rates[e].

Estimator	a[a]	Z[b]	p[c]
Combined[c]	0	-2.501	0.012
Combined	0.99	-2.150	0.032
Whole-clade[d]	0	-4.545	< 0.001
Whole-clade	0.99	-9.024	< 0.001

[a]Relative extinction rate, μ / λ.

[b]Z-transform test statistic (Whitlock 2005) from combining p-values across all lineages. For each lineage of a given standing diversity and stem age, we computed the one-tailed probability of species richness equal to or less than the observed standing diversity (Magallon & Sanderson 2001; eqn 10). We then combined these one-tailed probabilities using the unweighted Z-transform and doubled the resulting p-value, such that we were performing a two-tailed test for an excess of species-rich or species-poor lineages relative to the expectation under a homogeneous net diversification rate across all sphenomorphines.

[c]Combined taxonomic and phylogenetic estimator of net diversification rate described in this study.

[d]Estimator of net diversification rate from Magallon & Sanderson (2001), eqn 7.

[e]This test is conservative for several reasons. Strictly speaking, the test as performed is a two-tailed test for whether lineages are excessively species-poor or species rich, relative to the expected species diversity under the net diversification rate estimated

for the sphenomorphine clade as a whole. Thus, even if some lineages are excessively species-rich and others species-poor, these contrasting effects might cancel one another, such that no net effect is detected. This effect is present in the sphenomorphine tree, where most lineages are species-poor (figure 3), but two (*Ctenotus* and *Lerista*) are species-rich. Nonetheless, the signal is dominated by the 21 species-poor lineages, and we detect a significant tendency towards species-poor lineages (Table s4).

Another reason the test is conservative involves the fact that all lineages must contain at least a single species. This sets a lower bound on the minimum one-tailed probability for each lineage. For example, in the sphenomorphine data, *Eulamprus E* at 18.4 mya is the oldest sphenomorphine lineage with a single species in the present. However, under $a = 0$, and using the combined taxonomic-phylogenetic estimate of r described in this study, the probability of observing a level of diversity as extreme as this in the present is only $p = 0.085$. This is thus the minimum one-tailed probability that can be observed for any lineage with a single species in the present. This situation does not occur for lineages with excessive species diversity, where the probability of observing fewer than n lineages can be made arbitrarily close to 1 by taking n sufficiently large. This asymmetry (upper tail probabilities can be arbitrarily close to 1; lower-tail probabilities cannot be made arbitrarily close to 0) is a major weakness of this test.

An alternative approach might consider, for each lineage, the probability of observing a level of diversity as extreme or more extreme given the overall net diversification rate. These individual lineage tail probabilities could then be assessed using a Bonferroni-adjusted p-value. However, the asymmetry problem described above limits the utility of this test as well (in the example described above, it is clear that no lineage with a single species in the present can be significantly less diverse than expected under the homogeneous diversification model, since the minimum tail probability that can be observed is $p > 0.08$).

Despite the limitations of this test, *Ctenotus* is characterized by a significant excess of species under all values of a, even after Bonferroni-correction ($p < 0.0022$),

and *Lerista* shows significantly more species than expected for $a > 0.85$ ($p < 0.0022$). Of course, if we simply repeat all analyses by treating these sister taxa as a single lineage with 174 species, then the "*Lerista* + *Ctenotus*" clade is significantly more diverse than expected under the homogeneous diversification model across all values of a ($p < 0.002$).

Figure s1

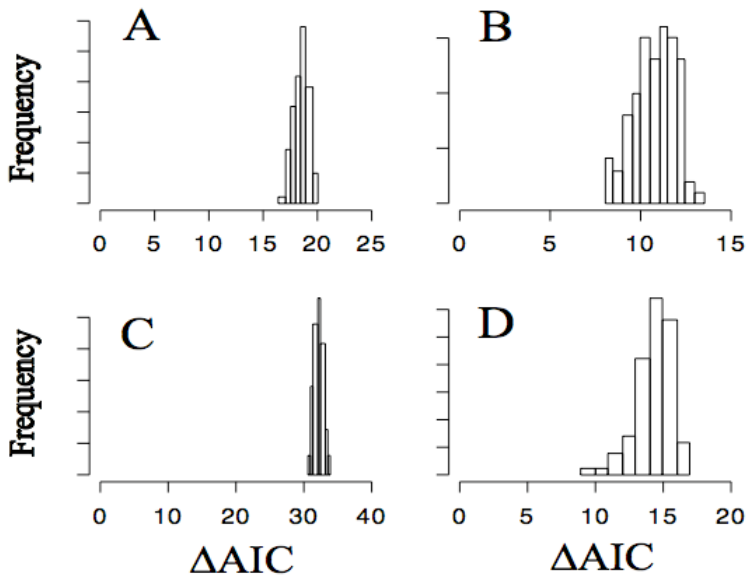

Figure s1. Summary of model-based analyses on 100 datasets generated by shuffling species diversities among major clades within paraphyletic genera (see text for details). Constant rate, flexible-rate, and rate-decrease models were fitted to each replicate dataset; in all 100 datasets (under both $a = 0$ and $a = 0.99$), the flexible-rate model fit the data significantly better than either alternative model and the reconstructed rate shift occurred in the MRCA of *Ctenotus* and *Lerista*. Histograms depict the difference in AIC scores (ΔAIC) between the two alternative models and the overall best fit model (flexible-rate) for replicate datasets: (A) difference between constant rate model and flexible-rate, $a = 0$; (B) rate-decrease and flexible-rate, $a = 0$; (C) constant rate and flexible-rate, $a = 0.99$; (D) rate-decrease and flexible-rate, $a = 0.99$. ΔAIC values are similar to those shown in Table 1. Results discussed in text are thus not conditional on the assignment of species diversity to paraphyletic genera depicted in figure 2.

Figure s2

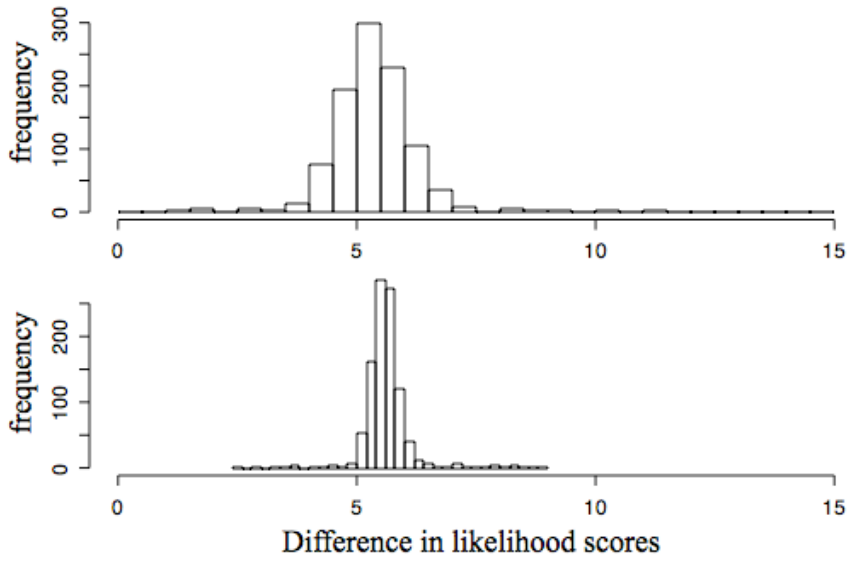

Figure s2. Posterior distribution of the difference in likelihood scores between rate-flexible and rate-decrease models under assumed relative extinction rates of $a = 0$ (upper) and $a = 0.99$ (lower). Rate-flexible model assumes that diversification rate r_1 shifts to r_2 at some topological location in the tree; rate-decrease model constrains the analysis such that *Ctenotus* retains the ancestral diversification rate present at the root node. Analyses based on 1000 PL trees sampled from the post-burnin distribution of topologies and branch lengths found through the Bayesian analysis. Likelihoods are always higher under the rate flexible model, which specified an increase in diversification rate for the node corresponding to the MRCA of the clade consisting exclusively of *Ctenotus* and *Lerista* in 97.7 % ($a = 0$) or 100% ($a = 0.99$) of sampled trees.

APPENDIX II

SUPPLEMENT TO CHAPTER 7

Supplementary Methods

Expected lineage-through-time curves: Because our results support density-dependent declines in diversification over an alternative model where rates vary continuously through time, we attempted to identify features of the expected lineage-through-time (LTT) curves that differed among models. There are two ways in which an expected LTT curve can differ from the observed curve: (1) the curves may differ qualitatively in shape, and (2) the curves may differ in the expected number of lineages produced after a fixed amount of time. We found the expected LTT curves under the maximum likelihood parameterizations of the model as

$$n(t) = 2\exp\left(\int_0^t \lambda(s)ds\right)$$

which gives the expected number of descendent lineages at any point in time *t*, starting with n = 2 initial lineages at time *t* = 0. It must be noted that the time-varying diversification models we employed assume fixed clade age and the models do not necessarily make sense if age is not limited to the value used to derive parameter estimates. For example, speciation rates for the continuous-decline model can become negative if t > K (eqn 2.3). There is thus no requirement that the maximum likelihood parameter estimates for a given diversification model will yield the observed number of lineages in the wood-warbler phylogeny. Expected LTT curves are shown in figure S2. In the case of the continuous-decline model, the maximum likelihood parameterization does not yield the correct number of lineages (n = 25) after 1.0 time units.

Simulations: To further explore the differences in fit between density-dependent and continuous decline models, we generated simulated LTT plots under the maximum likelihood parameter estimates for each model. As discussed above, the time-varying diversification models we employed require simulations of fixed clade age, as the parameters of the model specify a particular time course of change in the speciation rate and are conditional on the existence of an evolutionary process of the same age as the wood-warbler tree (1.0 time units). This implies that the number of lineages in existence at the end of each simulation is itself a random variable.

Most previous studies that have simulated time-varying diversification processes have used discrete-time phylogenetic simulation algorithms (e.g., Rabosky 2006b), in which phylogenetic trees are generated by iterating over a series of time steps such that each lineage has a probability of giving birth or going extinct each time step. However, this discrete-time approach merely approximates the continuous-time diversification process, and we used a simulation procedure that permits phylogenies to be simulated in continuous time (Rabosky & Lovette 2008).

We first divided the total simulation time (1.0 time units) into 100 intervals of $t = 0.01$ time units. We then calculated the value of the speciation rate (λ) for the midpoint of each interval (e.g., at $t = 0.01, 0.02, 0.03 \ldots 0.98, 0.99$) using the maximum likelihood parameter estimates for each model and equations 2.3. Each simulation was initiated with two lineages, which had parameter λ_1 on the first time interval; after $t = 0.01$ time units, parameters were updated to λ_2 and the simulation was continued to the end of the second time interval (overall elapsed time of 0.02 time units). These sequential parameter updates were continued until the end of the simulation. Thus, while we used a discrete approximation to model variation in λ, the underlying simulation occurred in continuous time.

In this approach, the number of lineages in the tree at the end of each simulation is itself a random variable. We wanted to compare simulated LTT plots for the set of

phylogenetic trees containing exactly 25 descendents at the end of the simulation (e.g., the same number of taxa as the wood-warbler tree). Under each simulation model, we simulated 500 trees of n = 25 taxa by retaining only those trees with exactly n = 25 taxa at the end of the simulation. Average LTT curves for each model are presented in figure S3 (a, b).

Simulated LTT curves resulting in n = 25 taxa at the end of the simulation look remarkably similar for density-dependent exponential and continuous-decline curves. However, this result must be considered in light of the fact that obtaining n = 25 taxa is much less likely under the maximum-likelihood parameter estimates for the continuous-decline model than the density-dependent model (e.g., figure S2). To demonstrate this directly, we simulated 10,000 phylogenies under maximum likelihood parameter estimates for each diversification model to look at the frequency distribution of clade size after 1.0 time units (figure S3, c, d).

Although the subset of LTT curves with n = 25 taxa appear similar among diversification models, the fitted models predict dramatically different species diversity (figure S3, c, d). Only the density-dependent model results in a pattern of expected species richness consistent with that observed in wood-warblers.

Figure S1

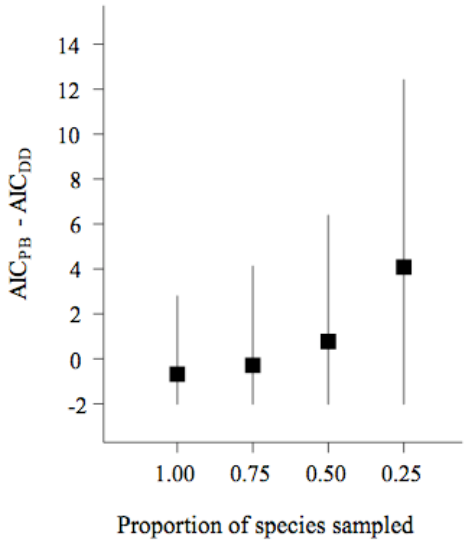

Figure S1. Tests for density-dependent diversification cannot simply contrast the likelihood of phylogenetic data under a constant-rate diversification model to a density dependent diversification model. To demonstrate this, we simulated sets of 5000 phylogenies under a constant speciation model with zero extinction, where phylogenies were 25%, 50%, 75%, or 100% complete at the species level [see Methods]. Each phylogeny was fitted with constant rate and density-dependent diversification models, and the difference in AIC scores tabulated as AIC_{PB} - AIC_{DD} , where AIC_{PB} and AIC_{DD} are AIC scores under pure-birth and best-fit density dependent diversification models. Shown are means and 95% confidence intervals around the distribution of this statistic. As the level of incomplete sampling increases, density dependent models provide a better fit to the pattern of lineage accumulation. This phenomenon occurs even though phylogenies were simulated under a constant speciation model and is driven solely by the fact that incomplete taxon sampling generates spurious temporal declines in diversification rates.

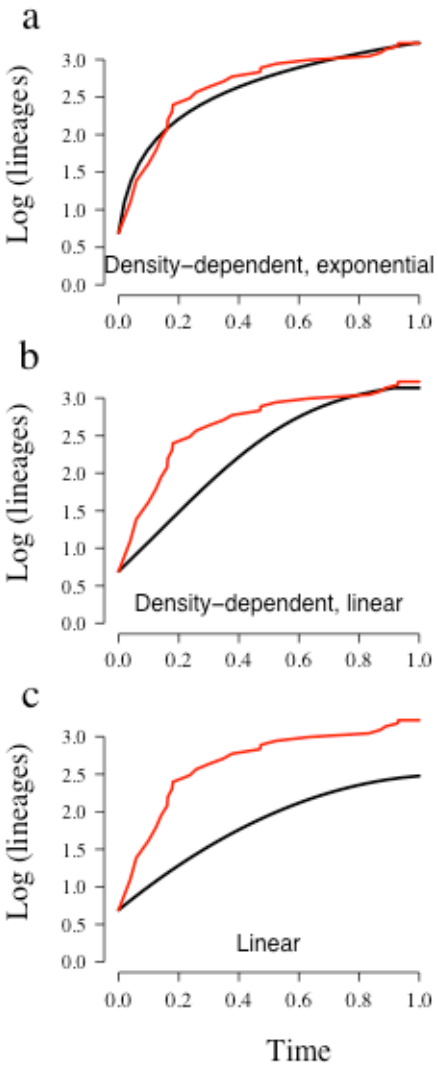

Figure S2. Expected lineage-through-time curves (black) under maximum likelihood parameterizations of the three rate-variable diversification models considered in this study: (a) density-dependent, exponential, (b) density-dependent, linear, and (c) continuous-decline (linear). Curves represent analytical expectations assuming maximum likelihood parameter estimates for each model. Red curve is observed lineage accumulation curve for wood-warblers (e.g., figure 2a). The maximum likelihood parameterization of the continuous-decline model fails to reach the observed species diversity for wood-warblers (n = 25).

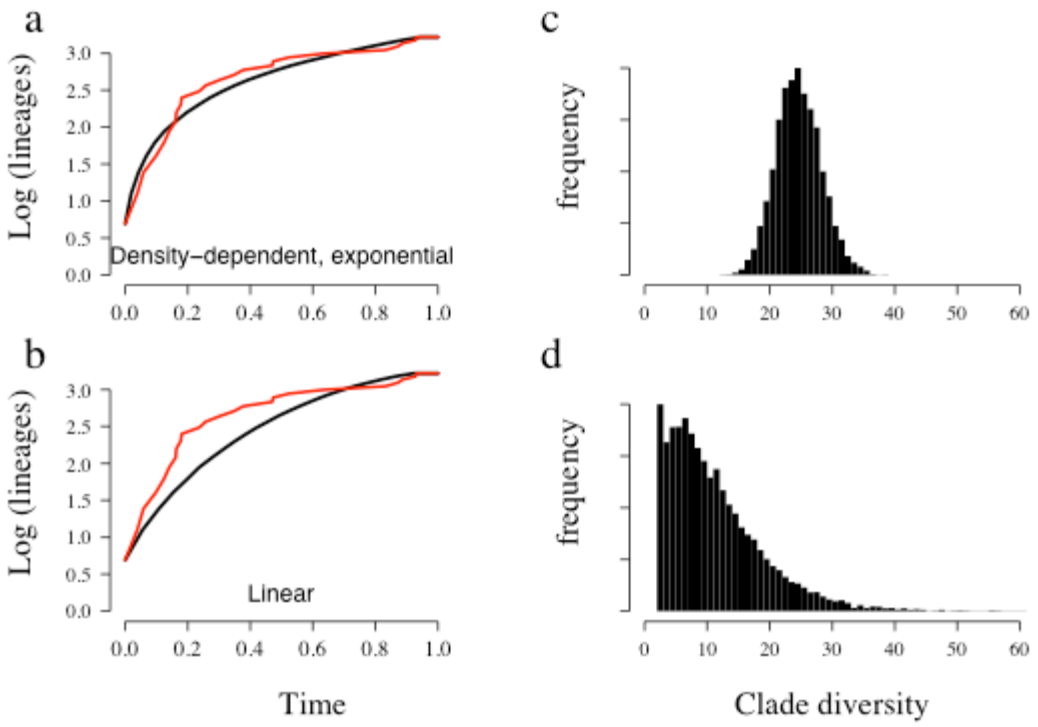

Figure S3. Expected lineage-through-time curves (black) for the subset of simulated phylogenies with exactly n = 25 lineages after 1.0 time units for density-dependent exponential (a) and continuous-decline (b) models. Phylogenies were simulated under maximum likelihood parameter estimates for each model. Although both of these curves appear to provide reasonable approximations of the observed lineage accumulation curve, this is in part illusory, as the maximum likelihood parameterization of the continuous-decline model is much less likely to result in the observed total number of lineages (n = 25) at the end of the simulation period. Frequency distributions of progeny lineages for phylogenetic trees simulated for 1.0 time units under these maximum likelihood parameters are shown for density-dependent exponential (c) and continuous-decline (d) models. Results are not shown for the density-dependent linear model, because the speciation rate becomes negative if the number of lineages in the simulation exceeds 25. Results for a and b based on 500 simulated phylogenies; results for c and d based on 10,000 simulated phylogenies.

APPENDIX III

SUPPLEMENT TO CHAPTER 8

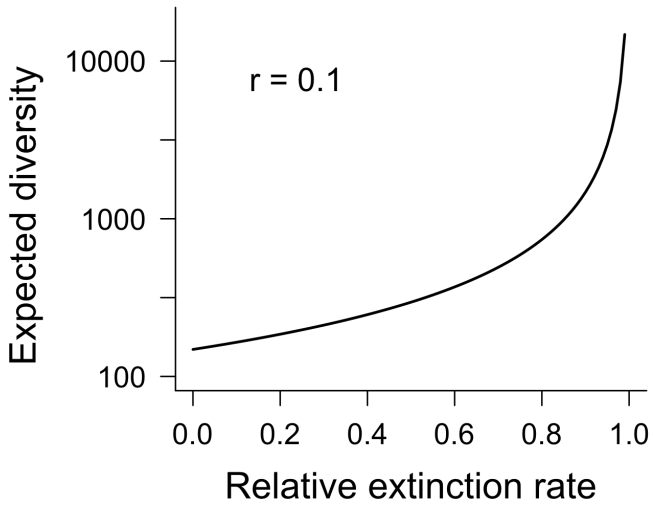

Figure S1. The expected extant clade diversity resulting from a birth-death process with a constant net rate of diversification (r) is dependent on the relative extinction rate (ε). Results shown are for clade diversifying for 50 my with a net diversification rate of r = 0.1 lineages/my and were calculated after Nee et al. (1994). This result pertains only to clades which have survived the present and are thus available to be observed.

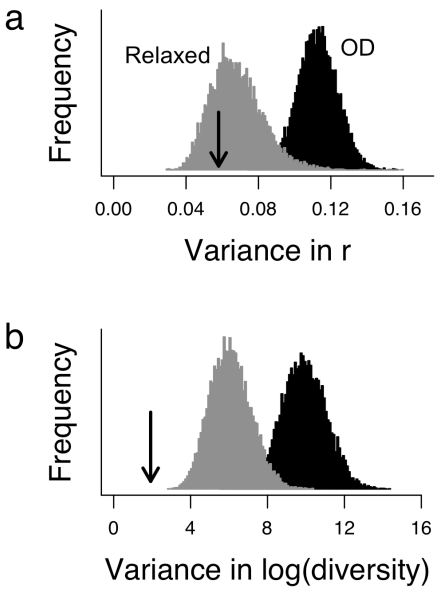

Figure S2. The variance in estimated net diversification rates (a) and species richness (b) for the avian tribes dataset (arrows) and for simulated rates and diversities under the relaxed rate model (grey histogram) and overdispersed rates model (black histogram). Rates under the relaxed model were drawn from a gamma distribution parameterized to fit the observed distribution of rates (see text for details); note that this results in a close relationship between the observed variance in r (arrow) and the simulated distribution. However, the resulting distribution of species richness has greater variance. The overdispersed model has much greater variance in both r and resulting species richness values. Results shown based on $\varepsilon = 0$, but results for $\varepsilon = 0.95$ and random ε are identical.

Table S1. Summary statistics for simulations described in text. ρ, observed correlations between age and diversity for the five datasets; Model, simulation model (CR = constant rate; RE = relaxed rate; OD = overdispersed); ρ_{SIM} , mean Pearson correlation between age and log(diversity) from simulations; sd, standard deviation of the simulated distribution; Δ, difference between observed and simulated correlations, scaled by the standard deviation { $\rho - \rho_{SIM}$) / sd}.

	Model	*ε*	ρ_{SIM}	*sd*	Δ	*p*
Angiosperms	CR	0	0.33	0.11	-4.16	<0.001
(ρ = -0.119)	CR	0.95	0.32	0.11	-4.03	<0.001
	CR	U	0.48	0.14	-4.35	<0.001
	REL	0	0.42	0.11	-4.8	<0.001
	RE	0.95	0.4	0.11	-4.48	<0.001
	RE	U	0.4	0.11	-4.5	<0.001
	OD	0	0.14	0.14	-1.84	0.033
	OD	0.95	0.13	0.14	-1.74	0.042
	OD	U	0.15	0.14	-1.84	0.034
Avian tribes	CR	0	0.69	0.06	-15.25	<0.001
(ρ = -0.188)	CR	0.95	0.4	0.09	-6.68	<0.001
	CR	U	0.54	0.08	-9.08	<0.001
	RE	0	0.34	0.09	-5.83	<0.001
	RE	0.95	0.23	0.1	-4	<0.001
	RE	U	0.18	0.1	-3.58	<0.001
	OD	0	0.35	0.09	-5.67	<0.001
	OD	0.95	0.3	0.1	-5.03	<0.001
	OD	U	0.32	0.1	-5.32	<0.001
Insects	CR	0	0.61	0.1	-0.76	0.226
(ρ = 0.523)	CR	0.95	0.78	0.07	-3.53	0.001

	CR	U	0.6	0.13	-0.43	0.323
	RE	0	0.6	0.12	-0.51	0.292
	RE	0.95	0.58	0.12	-0.32	0.35
	RE	U	0.57	0.13	-0.22	0.384
	OD	0	0.41	0.15	0.85	0.799
	OD	0.95	0.4	0.16	0.88	0.808
	OD	U	0.41	0.16	0.83	0.791
Mammals	CR	0	0.74	0.11	-6.54	<0.001
(ρ = 0.040)	CR	0.95	0.59	0.15	-3.69	0.001
	CR	U	0.6	0.15	-3.67	0.002
	RE	0	0.34	0.2	-1.48	0.077
	RE	0.95	0.22	0.22	-0.83	0.207
	RE	U	0.27	0.22	-1.04	0.147
	OD	0	0.34	0.21	-1.45	0.083
	OD	0.95	0.29	0.21	-1.16	0.13
	OD	U	0.32	0.21	-1.34	0.097
Teleosts	CR	0	0.9	0.03	-31.11	<0.001
(r = -0.074)	CR	0.95	0.84	0.05	-17.81	<0.001
	CR	U	0.82	0.06	-14.32	<0.001
	RE	0	0.46	0.15	-3.56	0.001
	RE	0.95	0.34	0.17	-2.49	0.011
	RE	U	0.39	0.16	-2.92	0.004
	OD	0	0.27	0.17	-2	0.029
	OD	0.95	0.27	0.17	-1.98	0.028
	OD	U	0.27	0.17	-2.03	0.026

REFERENCES

Adamowicz, S. J., A. Purvis, and M. A. Wills. 2008. Increasing morphological complexity in multiple parallel lineages of the Crustacea. *Proceedings of the National Academy of Sciences of the United States of America* 105:4786-4791.

Agapow, P.-M., and A. Purvis. 2002. Power of eight tree shape statistics to detect nonrandom diversification: a comparison by simulation of two models of cladogenesis. *Systematic Biology* 51:866-872.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267-281. *Second International Symposium on Information Theory*. Akademiai Kiado, Budapest.

Aldous, D., and L. Popovic. 2005. A critical branching process model for biodiversity. *Advances In Applied Probability* 37:1094-1115.

Aldous, D. J. 2001. Stochastic models and descriptive statistics for phylogenetic trees, from Yule to today. *Statistical Science* 16:23-34.

Alroy, J. 1996. Constant extinction, constrained diversification, and uncoordinated stasis in North American mammals. *Palaeogeography Palaeoclimatology Palaeoecology* 127:285-311.

Alroy, J. 2000. New methods for quantifying macroevolutionary patterns and processes. *Paleobiology* 26:707-733.

Alroy, J. 2008. The dynamics of origination and extinction in the marine fossil record. *Proceedings of the National Academy of Sciences of the U.S.A.* 105:11536-11542.

Alroy, J. 2009. Speciation and extinction in the fossil record of North American mammals. In R. Butlin, J. Bridle, and D. Schluter, eds. *Ecology and Speciation*. Cambridge Univ. Press, Cambridge.

Alroy, J., M. Aberhan, D. J. Bottjer, M. Foote, F. T. Fürsich, P. J. Harries, A. J. W. Hendy, S. M. Holland, L. C. Ivany, W. Kiessling, M. A. Kosnik, C. R. Marshall, A. J. McGowan, A. I. Miller, T. D. Olszewski, M. E. Patzkowsky, S. E. Peters, L. Villier, P. J. Wagner, N. Bonuso, P. S. Borkow, B. Brenneis, M. E. Clapham, L. M. Fall, C. A. Ferguson, V. L. Hanson, A. Z. Krug, K. M. Layou, E. H. Leckey, S. Nürnberg, C. M. Powers, J. A. Sessa, C. Simpson, A. Tomasovych, and C. C. Visaggi. 2008. Phanerozoic trends in the global diversity of marine invertebrates. *Science* 321:97-100.

Altekar, G., S. Dwarkadas, J. P. Huelsenbeck, and F. Ronquist. 2004. Parallel metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference. *Bioinformatics* 20:407-415.

Arbogast, B. S., S. V. Edwards, J. Wakeley, P. Beerli, and J. B. Slowinski. 2002. Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annual Review of Ecology and Systematics* 33:707-740.

Armbruster, W. S., M. E. Edwards, and E. M. Debevec. 1994. Floral character displacement generates assemblage structure of Western Australian triggerplants (Stylidium). *Ecology* 75:315-329.

Arnqvist, G., M. Edvardsson, U. Friberg, and T. Nilsson. 2000. Sexual conflict promotes speciation in insects. *Proceedings of the National Academy of Sciences of the United States of America* 97:10460-10464.

Avise, J. C. 2000. *Phylogeography.* Harvard University Press, Cambridge, MA.

Bailey, N. T. J. 1964. *The Elements of Stochastic Processes With Applications to the Natural Sciences*. Wiley, New York.

Baldwin, B. G., and M. J. Sanderson. 1998. Age and rate of diversification of the Hawaiian silversword alliance (Compositae). *Proceedings of the National Academy of Sciences of the U.S.A*. 95:9402-9406.

Barker, F. K., A. Cibois, P. Schikler, J. Feinstein, and J. Cracraft. 2004. Phylogeny and diversification of the largest avian radiation. *Proceedings of the National Academy of Sciences of the U.S.A* 101:11040-11045.

Barker, W. R., and P. J. M. Greenslade. 1982. *Evolution of the Flora and Fauna of Arid Australia*. Peacock Publications, Adelaide.

Barraclough, T. G. 2006. What can phylogenetics tell us about speciation in the Cape flora? *Diversity And Distributions* 12:21-26.

Barraclough, T. G., P. H. Harvey, and S. Nee. 1995. Sexual selection and taxonomic diversity in passerine birds. *Proceedings of the Royal Society of London Series B-Biological Sciences* 259:211-215.

Barraclough, T. G., J. E. Hogan, and A. P. Vogler. 1999. Testing whether ecological factors promote cladogenesis in a group of tiger beetles (Coleoptera : Cicindelidae). *Proceedings of the Royal Society of London Series B-Biological Sciences* 266:1061-1067.

Barraclough, T. G., and S. Nee. 2001. Phylogenetics and speciation. *Trends In Ecology & Evolution* 16:391-399.

Barraclough, T. G., and V. Savolainen. 2001. Evolutionary rates and species diversity in flowering plants. *Evolution* 55:677-683.

Barraclough, T. G., and A. P. Vogler. 2002. Recent diversification rates in North American tiger beetles estimated from a dated mtDNA phylogenetic tree. *Molecular Biology And Evolution* 19:1706-1716.

Benkman, C. W. 2003. Divergent selection drives the adaptive radiation of crossbills. *Evolution* 57:1176-1181.

Benton, M. J. 1995. Diversification and extinction in the history of life. *Science* 268:52-58.

Bernatchez, L., A. Chouinard, and G. Lu. 1999. Integrating molecular genetics and ecology in studies of adaptive radiation: whitefish, *Coregonus* sp., as a case study. *Biological Journal of the Linnean Society* 68:173-194.

Beuttell, K., and J. B. Losos. 1999. Ecological morphology of Caribbean anoles. *Herpetological Monographs*:1-28.

Bininda-Emonds, O. R. P., M. Cardillo, K. E. Jones, R. D. E. MacPhee, R. M. D. Beck, R. Grenyer, S. A. Price, R. A. Vos, J. L. Gittleman, and A. Purvis. 2007. The delayed rise of present-day mammals. *Nature* 446:507-512.

Bininda-Emonds, O. R. P., J. L. Gittleman, and M. A. Steel. 2002. The (Super)tree of life: Procedures, problems, and prospects. *Annual Review Of Ecology And Systematics* 33:265-289.

Bokma, F. 2008. Problems detecting density depedent diversification on phylogenies. Proc. R. Soc. Lond. B:10.1098/rspb.2008.1249.

Bolker, B. 2008. *Ecological Models and Data in R*. Princeton University Press, Princeton, NJ.

Bowler, J. M., ed. 1982. *Aridity in the late Tertiary and Quaternary of Australia.* Peacock Publications, Adelaide.

Brooks, S. P., and A. Gelman. 1997. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7:434-455.

Brown, J. M., M. A. McPeek, and M. L. May. 2000. A phylogenetic perspective on habitat shifts and diversity in the north American *Enallagma* damselflies. *Systematic Biology* 49:697-712.

Buckley, T. R., P. Arensburger, C. Simon, and G. K. Chambers. 2002. Combined data, Bayesian phylogenetics, and the origin of the New Zealand cicada genera. *Systematic Biology* 51:4-18.

Burnham, K. P., and D. R. Anderson. 2001. Kullback-Leibler information as a a basis for strong inference in ecological studies. *Wildlife Research* 28:111-119.

Burnham, K. P., and D. R. Anderson. 2002. *Model selection and multimodel inference - a practical information-theoretic approach*. Springer-Verlag, New York.

Byrd, R. H., P. Lu, J. Nocedal, and C. Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM J. Scientific Computing* 16:1190-1208.

Calsbeek, R., J. N. Thompson, and J. E. Richardson. 2003. Patterns of molecular evolution and diversification in a biodiversity hotspot: the California Floristic Province. *Molecular Ecology* 12:1021-1029.

Carr, T. R., and J. A. Kitchell. 1980. Dynamics of taxonomic diversity. *Paleobiology* 6:427-443.

Chan, K. M., and B. R. Moore. 2002. Whole-tree methods for detecting differential diversification rates. *Systematic Biology* 51:855-865.

Chan, K. M. A., and B. R. Moore. 2002. Whole-tree methods for detecting differential diversification rates. *Systematic Biology* 51:855-865.

Chapple, D. G., and J. S. Keogh. 2004. Parallel adaptive radiations in and and temperate Australia: molecular phylogeography and systematics of the Egernia whitii (Lacertilia : Scincidae) species group. *Biological Journal Of The Linnean Society* 83:157-173.

Cogger, H. G. 2000. *Reptiles and Amphibians of Australia*. Chelsea Green Publishing Co.

Cox, D. R., and D. Oakes. 1984. *Analysis of Survival Data*.

Coyne, J., and H. Orr. 2004. *Speciation*. Sinauer, Sunderland, MA.

Coyne, J. A., and H. A. Orr. 2004. *Speciation*. Sinauer, Cambridge.

Cracraft, J. 1982. A non-equilibrium theory for the rate-control of speciation and extinction and the origin of macroevolutionary patterns. *Systematic Zoology* 31:348-365.

Crisp, M., L. Cook, and D. Steane. 2004. Radiation of the Australian flora: what can comparisons of molecular phylogenies across multiple taxa tell us about the evolution of diversity in present-day communities? *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences* 359:1551-1571.

Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection*. John Murray, London.

Davies, T. J., T. G. Barraclough, M. W. Chase, P. S. Soltis, D. E. Soltis, and V. Savolainen. 2004. Darwin's abominable mystery: Insights from a supertree of the angiosperms. *Proceedings of the National Academy of Sciences of the United States of America* 101:1904-1909.

Davies, T. J., V. Savolainen, M. W. Chase, P. Goldblatt, and T. G. Barraclough. 2005. Environment, area, and diversification in the species-rich flowering plant family Iridaceae. *American Naturalist* 166:418-425.

Dodd, M. E., J. Silvertown, and M. W. Chase. 1999. Phylogenetic analysis of trait evolution and species diversity variation among angiosperm families. *Evolution* 53:732-744.

Drummond, A. J., S. Y. W. Ho, M. J. Phillips, and A. Rambaut. 2006. Relaxed phylogenetics and dating with confidence. *Plos Biology* 4:699-710.

Drummond, A. J., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 7.

Emerson, B. C., P. Oromi, and G. M. Hewitt. 2000. Tracking colonization and diversification of insect lineages on islands: mitochondrial DNA

phylogeography of *Tarphius canariensis* (Coleoptera : Colydiidae) on the Canary Islands. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 267:2199-2205.

Emerson, B. C., P. Oromi, and G. M. Hewitt. 2000. Colonization and diversification of the species Brachyderes rugatus (Coleoptera) on the Canary Islands: Evidence from mitochondrial DNA COII gene sequences. *Evolution* 54:911-923.

Emerson, B. C., E. Paradis, and C. Thebaud. 2001. Revealing the demographic histories of species using DNA sequences. *Trends In Ecology & Evolution* 16:707-716.

Erwin, D. H., and J. W. Valentine. 1984. Hopeful monsters, transposons, and the metazoan radiation. *Proceedings of the National Academy of Sciences of the U.S.A*. 81.

Erwin, D. H., J. W. Valentine, and J. J. Sepkoski. 1987. A comparative-study of diversification events - the early paleozoic versus the mesozoic. *Evolution* 41:1177-1186.

Evans, K. L., and K. J. Gaston. 2005. Can the evolutionary-rates hypothesis explain species-energy relationships? *Functional Ecology* 19:899-915.

Farrell, B. D. 1998. "Inordinate fondness" explained: Why are there so many beetles? *Science* 281:555-559.

Felsenstein, J. 1985. Phylogenies and the comparative method. *American Naturalist* 125:1-15.

Felsenstein, J. 2004. *Inferring Phylogenies*. Sinauer: Sunderland, MA.

Fine, P. V. A., and R. H. Ree. 2006. Evidence for a time-integrated species-area effect on the latitudinal gradient in tree diversity. *American Naturalist* 168:796-804.

Foote, M. 1996. Ecological controls on the evolutionary recovery of post-Paleozoic crinoids. *Science* 274:1492-1495.

Freckleton, R. P., A. B. Phillimore, and M. Pagel. 2008. Relating traits to diversification: A simple test. *American Naturalist* 172:102-115.

Friesen, V. L., B. C. Congdon, H. E. Walsh, and T. P. Birt. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. *Molecular Ecology* 6:1047-1058.

Fryer, G., and T. D. Iles. 1972. *The Cichlid Fishes of the Great Lakes of Africa*. Oliver and Boyd, Edinburgh.

Fulton, M., and S. A. Hodges. 1999. Floral isolation between Aquilegia formosa and Aquilegia pubescens. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 266:2247-2252.

Gavrilets, S., and A. Vose. 2005. Dynamic patterns of adaptive radiation. *Proc. Nat. Acad. Sci. U.S.A.* 102:18040-18045.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* 7:457-511.

Giannasi, N., R. S. Thorpe, and A. Malhotra. 2000. A phylogenetic analysis of body size evolution in the *Anolis roquet* group (Sauria : Iguanidae): character displacement or size assortment? *Molecular Ecology* 9:193-202.

Giannini, N. P. 2003. Canonical phylogenetic ordination. *Systematic Biology* 52:684-695.

Gianoli, E. 2004. Evolution of a climbing habit promotes diversification in flowering plants. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271:2011-2015.

Gilinsky, N. L. 1994. Volatility and the Phanerozoic decline of background extinction intensity. *Paleobiology* 20:445-458.

Gillespie, R. 2004. Community assembly through adaptive radiation in Hawaiian spiders. *Science* 303:356-359.

Gittenberger, E. 1991. What about non-adaptive radiation? *Biological Journal Of The Linnean Society* 43:263-272.

Givnish, T. J., ed. 1997. *Adaptive radiation and molecular systematics: issues and approaches.* Cambridge University Press, New York.

Givnish, T. J., and K. J. Sytsma. 1997. Consistency, characters, and the likelihood of correct phylogenetic inference. *Molecular Phylogenetics and Evolution* 7:320-330.

Givnish, T. J., and K. J. Sytsma, eds. 1997. *Molecular Evolution and Adaptive Radiation*. Cambridge University Press, New York.

Glor, R. E., J. J. Kolbe, R. Powell, A. Larson, and J. B. Losos. 2003. Phylogenetic analysis of ecological and morphological diversification in Hispaniolan trunk-ground anoles (*Anolis cybotes* group). *Evolution* 57:2383-2397.

Goldberg, E. E., K. Roy, R. Lande, and D. Jablonski. 2005. Diversity, endemism, and age distributions in macroevolutionary sources and sinks. *American Naturalist* 165:623-633.

Good-Avila, S. V., V. Souza, B. S. Gaut, and L. E. Eguiarte. 2006. Timing and rate of speciation in Agave (Agavaceae). *Proceedings Of The National Academy Of Sciences of the United States of America* 103:9124-9129.

Grant, P. R. 1986. *Ecology and Evolution of Darwin's Finches*. Princeton University Press, Princeton.

Grant, P. R., and B. R. Grant. 2007. *How and Why Species Multiply*. Princeton Univ. Press, Princeton.

Grantham, T. A. 1995. Hierarchical approaches to macroevolution - recent work on species selection and the effect hypothesis. *Annual Review of Ecology and Systematics* 26:301-321.

Greenwood, P. 1984. African cichlids and evolutionary theories in A. Echelle, and I. Kornfield, eds. *Evolution of Fish Species Flocks*. Univ. of Maine Press, Orono.

Gurevitch, J., P. S. Curtis, and M. H. Jones. 2001. Meta-analysis in ecology. Pp. 199-247. *Advances In Ecological Research,* Vol 32.

Hansen, T. 1980. Influence of larval dispersal and geographic distribution on species longevity in neogastropods. *Paleobiology* 6:193-207.

Harmon, L. J., J. A. Schulte, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation in iguanian lizards. *Science* 301:961-964.

Harmon, L. J., J. T. Weir, C. Brock, R. E. Glor, and W. E. Challenger. 2008. GEIGER: Investigating evolutionary radiations. *Bioinformatics* 24.

Hawkins, J. A. 2006. Using phylogeny to investigate the origins of the Cape flora: the importance of taxonomic, gene and genome sampling strategies. *Diversity and Distributions* 12:27-33.

Heard, S. B., and A. Mooers. 2000. Measuring the loss of evolutionary history from extinction: phylogenetically patterned speciation rates and extinction risks alter the calculus of biodiversity. *Proc. R. Soc. Lond. B* 267:613-620.

Hey, J. 1992. Using phylogenetic trees to study speciation and extinction. *Evolution* 46:627-640.

Hodges, S. A. 1997. Floral nectar spurs and diversification. *Int. J. Plant. Sci*. 158:S81-S88.

Hodges, S. A., and M. L. Arnold. 1995. Spurring plant diversification: Are floral nectar spurs a key innovation? *Proceedings of the Royal Society of London Series B-Biological Sciences* 262:343-348.

Hubbell, S. P. 1979. Tree Dispersion, Abundance, and Diversity in a Tropical Dry Forest. Science 203:1299-1309.

Hubbell, S. P. 2000. *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton, N.J.

Huelsenbeck, J. P., and K. A. Crandall. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review Of Ecology And Systematics* 38:437-466.

Huelsenbeck, J. P., and F. Ronquist. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.

Hunt, T., J. Bergsten, Z. Levkanicova, A. Papadopoulou, O. S. John, R. Wild, P. M. Hammond, D. Ahrens, M. Balke, M. S. Caterino, J. Gomez-Zurita, I. Ribera, T. G. Barraclough, M. Bocakova, L. Bocak, and A. P. Vogler. 2007. A comprehensive phylogeny of beetles reveals the evolutionary origins of a superradiation. *Science* 318:1913-1916.

Hutchinson, G. E. 1959. Homage to Santa Rosalia. *American Naturalist* 93:145-159.

Isaac, N. J. B., K. E. Jones, J. L. Gittleman, and A. Purvis. 2005. Correlates of species richness in mammals: Body size, life history, and ecology. *American Naturalist* 165:600-607.

Jablonski, D. 1986. Larval ecology and macroevolution in marine-invertebrates. *Bulletin of Marine Science* 39:565-587.

Jablonski, D. 2000. Micro- and macroevolution: scale and hierarchy in evolutionary biology and paleobiology. *Paleobiology* 26:15-52.

Jablonski, D. 2008. Species Selection: Theory and Data. *Annual Review of Ecology Evolution and Systematics* 39:501-524.

James, C. D., and R. Shine. 2000. Why are there so many coexisting species of lizards in Australian deserts? *Oecologia* 125:127-141.

Jaramillo, C., M. J. Rueda, and G. Mora. 2006. Cenozoic plant diversity in the Neotropics. *Science* 311:1893-1896.

Jeffries, M. J., and J. H. Lawton. 1984. Enemy Free Space And The Structure Of Ecological Communities. *Biological Journal Of The Linnean Society* 23:269-286.

Johns, G. C., and J. C. Avise. 1998. Tests for ancient species flocks based on molecular phylogenetic appraisals of *Sebastes* rockfishes and other marine fishes. *Evolution* 52:1135-1146.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. *Trends In Ecology & Evolution* 19:101-108.

Jones, K. E., O. R. P. Bininda-Emonds, and J. L. Gittleman. 2005. Bats, clocks, and rocks: Diversification patterns in chiroptera. *Evolution* 59:2243-2255.

Jones, K. E., A. Purvis, A. MacLarnon, O. R. P. Bininda-Emonds, and N. B. Simmons. 2002. A phylogenetic supertree of the bats (Mammalia : Chiroptera). *Biological Reviews* 77:223-259.

Kadereit, J. W., E. M. Griebeler, and H. P. Comes. 2004. Quaternary diversification in European alpine plants: pattern and process. *Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences* 359:265-274.

Kalisz, S., R. H. Ree, and R. D. Sargent. 2006. Linking floral symmetry genes to breeding system evolution. *Trends in Plant Science* 11:568-573.

Katzourakis, A., A. Purvis, S. Azmeh, G. Rotheray, and F. Gilbert. 2001. Macroevolution of hoverflies (Diptera : Syrphidae): the effect of using higher-level taxa in studies of biodiversity, and correlates of species richness. *Journal of Evolutionary Biology* 14:219-227.

Katzourakis, A., A. Rambaut, and O. G. Pybus. 2005. The evolutionary dynamics of endogenous retroviruses. *Trends In Microbiology* 13:463-468.

Kay, K., C. Voelckel, J. Yan, K. Hufford, D. Kaska, and S. A. Hodges. 2006. Floral characters and species diversification in L. Harder, and S. Barrett, eds. *Ecology and Evolution of Flowers.* Oxford Univ. Press, Oxford.

Kendall, D. G. 1948. On the generalized "birth-and-death" process. *Ann. Math. Stat*. 19:1-15.

Kirschner, M., and J. Gerhart. 1998. Evolvability. *Proceedings of the National Academy of Sciences of the United States of America* 95:8420-8427.

Klak, C., G. Reeves, and T. Hedderson. 2004. Unmatched tempo of evolution in Southern African semi-desert ice plants. *Nature* 427:63-65.

Kozak, K. H., D. W. Weisrock, and A. Larson. 2006. Rapid lineage accumulation in a non-adaptive radiation: phylogenetic analysis of diversification rates in eastern North American woodland salamanders (Plethodontidae : Plethodon). *Proceedings Of The Royal Society B-Biological Sciences* 273:539-546.

Kruger, O. 2008. Engines of speciation: a comparative study in birds of prey. *Journal of Evolutionary Biology* 21:861-872.

Kubo, T., and Y. Iwasa. 1995. Inferring the rates of branching and extinction from molecular phylogenies. *Evolution* 49:694-704.

Labandeira, C. C., and J. J. Sepkoski. 1993. Insect Diversity In The Fossil Record. *Science* 261:310-315.

Lack, D. 1947. *Darwin's Finches*. Cambridge University Press, Cambridge.

Levinton, J. S. 1979. A theory of diversity equilibrium and morphological evolution. *Science* 204:335-336.

Lieberman, B. S. 2001. A test of whether rates of speciation were unusually high during the Cambrian radiation. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 268:1707-1714.

Linder, H. P., P. Eldenas, and B. G. Briggs. 2003. Contrasting patterns of radiation in African and Australian Restionaceae. *Evolution* 57:2688-2702.

Liow, L. H., M. Fortelius, E. Bingham, K. Lintulaakso, H. Mannila, L. Flynn, and N. C. Stenseth. 2008. Higher origination and extinction rates in larger mammals. *Proc. Nat. Acad. Sci. U.S.A.* 105:6097-6102.

Lloyd, E. A., and S. J. Gould. 1993. Species selection on variability. *Proceedings of the National Academy of Sciences of the United States of America* 90:595-599.

Losos, J. B., and D. Schluter. 2000. Analysis of an evolutionary species-area relationship. *Nature* 408:847-849.

Lovette, I. J., and E. Bermingham. 1999. Explosive speciation in the New World Dendroica warblers. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 266:1629-1636.

Lovette, I. J., and E. Bermingham. 2001. Mitochondrial perspective on the phylogenetic relationships of the Parula wood-warblers. *Auk* 118:211-215.

Lovette, I. J., and E. Bermingham. 2002. What is a wood-warbler? Molecular characterization of a monophyletic parulidae. *Auk* 119:695-714.

Lovette, I. J., E. Bermingham, and R. E. Ricklefs. 2002. Clade-specific morphological diversification and adaptive radiation in Hawaiian songbirds. *Proceedings of the Royal Society of London Series B-Biological Sciences* 269:37-42.

Lovette, I. J., and W. M. Hochachka. 2006. Simultaneous effects of phylogenetic niche conservatism and competition on avian community structure. *Ecology* 87:S14-S28.

Macarthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous forests. *Ecology* 39:599-619.

MacArthur, R. H. 1969. Patterns of communities in the tropics. *Biological Journal of the Linnean Society* 1:19-30.

MacArthur, R. H., and E. O. Wilson. 1963. An equilibrium theory of insular zoogeography. *Evolution* 17:373-387.

Macey, J. R., J. A. Schulte, A. Larson, and T. J. Papenfuss. 1998. Tandem duplication via light-strand synthesis may provide a precursor for mitochondrial genomic rearrangement. *Molecular Biology And Evolution* 15:71-75.

Maddison, W. P. 2006. Confounding asymmetries in evolutionary diversification and character change. *Evolution* 60:1743-1746.

Maddison, W. P., P. E. Midford, and S. P. Otto. 2007. Estimating a binary character's effect on speciation and extinction. *Systematic Biology* 56:701-710.

Magallon, S., and M. J. Sanderson. 2001. Absolute diversification rates in angiosperm clades. *Evolution* 55:1762-1780.

Mank, J. E. 2007. Mating preferences, sexual selection and patterns of cladogenesis in ray-finned fishes. *Journal of Evolutionary Biology* 20:597-602.

Marjoram, P., J. Molitor, V. Plagnol, and S. Tavare. 2003. Markov chain Monte Carlo without likelihoods. *Proc. Nat. Acad. Sci. U.S.A.* 100:15324 –15328.

Markgraf, V., M. McGlone, and G. Hope. 1995. Neogene paleoenvironmental and paleoclimatic change in southern temperate ecosystems - a southern perspective. *Trends. Ecol. Evol.* 10:143-147.

Martin, A. P., E. K. Costello, A. F. Meyer, D. R. Nemergut, and S. K. Schmidt. 2004. The rate and pattern of cladogenesis in microbes. *Evolution* 58:946-955.

Martin, P. R., and T. E. Martin. 2001. Ecological and fitness consequences of species coexistence: A removal experiment with wood warblers. *Ecology* 82:189-206.

McConway, K. J., and H. J. Sims. 2004. A likelihood-based method for testing for nonstochastic variation of diversification rates in phylogenies. *Evolution* 58:12-23.

McPeek, M. A. 2008. Ecological dynamics of clade diversification and community assembly. *American Naturalist* 172.

McPeek, M. A., and J. M. Brown. 2000. Building a regional species pool: Diversification of the Enallagma damselflies in eastern North America. *Ecology* 81:904-920.

McPeek, M. A., and J. M. Brown. 2007. Clade age and not diversification rate explains species richness among animal taxa. *American Naturalist* 169:E97-E106.

Mitter, C., B. Farrell, and B. Wiegmann. 1988. The phylogenetic study of adaptive zones - has phytophagy promoted insect diversification*? American Naturalist* 132:107-128.

Mooers, A. O., and S. B. Heard. 1997. Evolutionary process from phylogenetic tree shape. *Quarterly Review of Biology* 72:31-54.

Moore, B. R., K. M. Chan, and M. J. Donoghue. 2004. Detecting diversification rate variation in supertrees. Pp. 487-533 in O. P. Bininda-Emonds, ed. *Phylogenetic supertrees: combining information to reveal the tree of life.* Kluwer Academic, Dordrecht.

Moore, B. R., and M. J. Donoghue. 2007. Correlates of diversification in the plant clade dipsacales: Geographic movement and evolutionary innovations. *American Naturalist* 170:S28-S55.

Morrow, E. H., T. E. Pitcher, and G. Arnqvist. 2003. No evidence that sexual selection is an 'engine of speciation' in birds. *Ecology Letters* 6:228-234.

Morse, D. H. 1989. *American Warblers: An Ecological and Behavioral Perspective*. Harvard University Press, Cambridge, MA.

Morton, S. R., and C. D. James. 1988. The diversity and abundance of lizards in arid Australia - a new hypothesis. *American Naturalist* 132:237-256.

Near, T. J., T. W. Kassler, J. B. Koppelman, C. B. Dillman, and D. P. Philipp. 2003. Speciation in North American black basses, Micropterus (Actinopterygii : Centrarchidae). *Evolution* 57:1610-1621.

Nee, S. 2001. Inferring speciation rates from phylogenies. *Evolution* 55:661-668.

Nee, S. 2006. Birth-death models in macroevolution. *Ann. Rev. Ecol. Evol. Syst*. 37:1-17.

Nee, S., E. C. Holmes, R. M. May, and P. H. Harvey. 1994. Extinction rates can be estimated from molecular phylogenies. *Proc. R. Soc. Lond. B.* 344:77-82.

Nee, S., R. M. May, and P. H. Harvey. 1994. The reconstructed evolutionary process. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* 344:305-311.

Nee, S., A. Mooers, and P. H. Harvey. 1992. Tempo and mode of evolution revealed from molecular phylogenies. *Proceedings of the National Academy of Sciences of the U.S.A.* 89:8322-8326.

Nee, S., A. O. Mooers, and P. H. Harvey. 1992. Tempo and Mode of Evolution Revealed from Molecular Phylogenies. *Proceedings of the National Academy of Sciences of the U.S.A*. 89:8322-8326.

Newman, M. E. J., and P. Sibani. 1999. Extinction, diversity and survivorship of taxa in the fossil record. *Proceedings of the Royal Society Of London Series B-Biological Sciences* 266:1593-1599.

Nicolakakis, N., D. Sol, and L. Lefebvre. 2003. Behavioural flexibility predicts species richness in birds, but not extinction risk. *Animal Behaviour* 65:445-452.

Nunn, C. L., S. Altizer, W. Sechrest, K. E. Jones, R. A. Barton, and J. L. Gittleman. 2004. Parasites and the evolutionary diversification of primate clades. *American Naturalist* 164:S90-S103.

Nylander, J. S. 2004. MrModeltest. URL http://www.abc.se/wnylander/.

Okasha, S. 2006. *Evolution and the Levels of Selection*. Oxford Univ Press, Oxford.

Osenberg, C. W., O. Sarnelle, and D. E. Goldberg. 1999. Meta-analysis in ecology: Concepts, statistics, and applications. *Ecology* 80:1103-1104.

Owens, I. P. F., P. M. Bennett, and P. H. Harvey. 1999. Species richness among birds: body size, life history, sexual selection or ecology? *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 266:933-939.

Pagel, M. D. 1992. A Method for the Analysis of Comparative Data. *Journal Of Theoretical Biology* 156:431-442.

Paradis, E. 1997. Assessing temporal variations in diversification rates from phylogenies: estimation and hypothesis testing. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 264:1141-1147.

Paradis, E. 1998. Detecting shifts in diversification rates without fossils. *American Naturalist* 152:176-187.

Paradis, E. 1998. Testing for constant diversification rates using molecular phylogenies: A general approach based on statistical tests for goodness of fit. *Molecular Biology And Evolution* 15:476-479.

Paradis, E. 2003. Analysis of diversification: combining phylogenetic and taxonomic data. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 270:2499-2505.

Paradis, E. 2004. Can extinction rates be estimated without fossils? *Journal Of Theoretical Biology* 229:19-30.

Paradis, E. 2005. Statistical analysis of diversification with species traits. *Evolution* 59:1-12.

Paradis, E. 2008. Asymmetries in phylogenetic diversification and character change can be untangled. *Evolution* 62:241-247.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20:289-290.

Parmakelis, A., M. Pfenninger, L. Spanos, G. Papagiannakis, C. Louis, and M. Mylonas. 2005. Inference of a radiation in Mastus (gastropoda, pulmonata, enidae) on the island of Crete. *Evolution* 59:991-1005.

Pearson, P. N. 1996. Cladogenetic, extinction and survivorship patterns from a lineage phylogeny: the Paleogene planktonic foraminifera. *Micropaleontology* 42:179-188.

Phillimore, A. B., R. P. Freckleton, C. D. L. Orme, and I. P. F. Owens. 2006. Ecology predicts large-scale patterns of phylogenetic diversification in birds. *American Naturalist* 168:220-229.

Phillimore, A. B., and T. D. Price. 2008. Density dependent cladogenesis in birds. *Plos Biology* 6:e71.

Pianka, E. R. 1972. Zoogeography and Speciation of Australian Desert Lizards - Ecological Perspective. *Copeia* 1972:127

Pianka, E. R. 1986. *Ecology and Natural History of Desert Lizards*. Princeton University Press, Princeton, N.J.

Piessens, R., d.-K. E., C. Uberhuber, and D. Kahaner. 1983. *Quadpack: a Subroutine Package for Automatic Integration*. Springer-Verlag.

Pigliucci, M. 2008. Opinion - Is evolvability evolvable? *Nature Reviews Genetics* 9:75-82.

Pitra, C., J. Fickel, E. Meijaard, and P. C. Groves. 2004. Evolution and phylogeny of old world deer. *Molecular Phylogenetics And Evolution* 33:880-895.

Plagnol, V., and S. Tavare. 2004. Approximate Bayesian computation and MCMC. Pp. 99-114 in N. H, ed. *Monte Carlo and Quasi-Monte Carlo Methods*. Springer-Verlag.

Pol, D. 2004. Empirical problems of the hierarchical likelihood ratio test for model selection. *Systematic Biology* 53:949-962.

Posada, D., and T. R. Buckley. 2004. Model selection and model averaging approaches in phylogenetics: advantages of the Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Systematic Biology* 53:793-808.

Posada, D., and K. A. Crandall. 2001. Selecting the best-fit model of nucleotide substitution. *Systematic Biology* 50:580-601.

Price, T. 2008. *Speciation in Birds*. Roberts and Company, Greenwood Village, CO.

Price, T., I. J. Lovette, E. Bermingham, H. L. Gibbs, and A. D. Richman. 2000. The imprint of history on communities of North American and Asian warblers. *American Naturalist* 156:354-367.

Purvis, A. 2008. Phylogenetic approaches to the study of extinction. *Ann. Rev. Ecol. Evol. Syst.* 39:301-319.

Purvis, A., P.-M. Agapow, J. L. Gittleman, and G. M. Mace. 2000. Nonrandom extinction risk and the loss of evolutionary history. *Science* 288:328-330.

Purvis, A., and P. M. Agapow. 2002. Phylogeny imbalance: Taxonomic level matters. *Systematic Biology* 51:844-854.

Purvis, A., M. Cardillo, R. Grenyer, and B. Collen. 2005. Correlates of extinction risk: phylogeny, biology, threat and scale. in A. Purvis, A. Brooks, and J. Gittleman, eds. *Phylogeny and Conservation*. Cambridge Univ. Press, Cambridge.

Purvis, A., C. D. L. Orme, N. H. Toomey, and P. N. Pearson. 2009. Temporal patterns in diversification rates in R. Butlin, J. Bridle, and D. Schluter, eds. *Speciation and Patterns of Diversity*. Cambridge Univ. Press, Cambridge.

Pybus, O. G., and P. H. Harvey. 2000. Testing macro-evolutionary models using incomplete molecular phylogenies. *Proceedings Of The Royal Society Of London Series B-Biological Sciences* 267:2267-2272.

Rabosky, D. L. 2006. Likelihood methods for detecting temporal shifts in diversification rates. *Evolution* 60:1152-1164.

Rabosky, D. L. 2006. LASER: a maximum likelihood toolkit for detecting temporal shifts in diversification rates. *Evolutionary Bioinformatics Online* 2:257-260.

Rabosky, D. L. 2009. Ecological limits on clade diversification in higher taxa. *American Naturalist* In press.

Rabosky, D. L., S. C. Donnellan, A. L. Talaba, and I. J. Lovette. 2007. Exceptional among-lineage variation in diversification rates during the radiation of Australia's most diverse vertebrate clade. *Proceedings of the Royal Society B-Biological Sciences* 274:2915-2923.

Rabosky, D. L., and I. J. Lovette. 2008. Explosive evolutionary radiations: Decreasing speciation or increasing extinction through time? *Evolution* 62:1866-1875.

Rabosky, D. L., and I. J. Lovette. 2008. Density dependent diversification in North American wood warblers. *Proceedings of the Royal Society B-Biological Sciences* 275: 2363-2371

Rabosky, D. L., and I. J. Lovette. 2008. Problems detecting density-dependent diversification on phylogenies: reply to Bokma. *Proceedings of the Royal Society B-Biological Sciences* doi: 10.1098/rspb.2008.1158.

Rabosky, D. L., J. Reid, M. A. Cowan, and J. Foulkes. 2007. Community-wide overdispersion of body size in Australian desert lizard communities. *Oecologia* 154:561-570.

Rabosky, D. L., and U. Sorhannus. 2009. Diversity dynamics of marine planktonic diatoms across the Cenozoic. *Nature* 457:183-186.

Rankin, D., and A. Lopez-Sepulcre. 2005. Can adaptation lead to extinction? *Oikos* 111.

Raup, D. M. 1985. Mathematical-Models Of Cladogenesis. *Paleobiology* 11:42-52.

Raup, D. M., S. J. Gould, T. J. M. Schopf, and Simberlo.Ds. 1973. Stochastic models of phylogeny and evolution of diversity. *Journal of Geology* 81:525-542.

Rawlings, L. H., D. L. Rabosky, S. C. Donnellan, and M. N. Hutchinson. 2008. Python phylogenetics: inferences from morphology and mitochondrial DNA. *Biol. J. Linn. Soc*. 93:603-619.

Ree, R. H. 2005. Detecting the historical signature of key innovations using stochastic models of character evolution and cladogenesis. *Evolution* 59:257-265.

Reeder, T. W. 2003. A phylogeny of the Australian Sphenomorphus group (Scincidae : Squamata) and the phylogenetic placement of the crocodile skinks (Tribolonotus): Bayesian approaches to assessing congruence and obtaining confidence in maximum likelihood inferred relationships. *Molecular Phylogenetics and Evolution* 27:384-397.

Revell, L. J., L. J. Harmon, and R. E. Glor. 2005. Underparameterized model of sequence evolution leads to bias in the estimation of diversification rates from molecular phylogenies. *Systematic Biology*. 54:973-983.

Ribera, I., T. G. Barraclough, and A. P. Vogler. 2001. The effect of habitat type on speciation rates and range movements in aquatic beetles: inferences from species-level phylogenies. *Molecular Ecology* 10:721-735.

Rice, W. R., and E. E. Hostert. 1993. Laboratory experiments on speciation: what have we learned in 40 years? *Evolution* 47:1637-1653.

Richmond, J. Q., and E. L. Jockusch. 2007. Body size evolution simultaneously creates and collapses species boundaries in a clade of scincid lizards. *Proceedings Of The Royal Society B-Biological Sciences* 274:1701-1708.

Ricklefs, R. E. 2003. Global diversification rates of passerine birds. *Proceedings of the Royal Society of London Series B-Biological Sciences* 270:2285-2291.

Ricklefs, R. E. 2006. Global variation in the diversification rate of passerine birds. *Ecology* 87:2468-2478.

Ricklefs, R. E. 2007. Estimating diversification rates from phylogenetic information. *Trends in Ecology & Evolution* 22:601-610.

Ricklefs, R. E. 2008. Speciation, extinction, and diversity. *Ecological Reviews* In press.

Ricklefs, R. E., and G. W. Cox. 1972. Taxon cycles and the West Indian avifauna. *American Naturalist* 106:195-219.

Ricklefs, R. E., J. B. Losos, and T. M. Townsend. 2007. Evolutionary diversification of clades of squamate reptiles. *Journal of Evolutionary Biology* 20:1751-1762.

Ricklefs, R. E., and I. J. Lovette. 1999. The roles of island area per se and habitat diversity in the species-area relationships of four Lesser Antillean faunal groups. *Journal of Animal Ecology* 68:1142-1160.

Ricklefs, R. E., A. E. Schwarzbach, and S. S. Renner. 2006. Rate of lineage origin explains the diversity anomaly in the world's mangrove vegetation. *American Naturalist* 168:805-810.

Ritchie, M. G., S. A. Webb, J. A. Graves, A. E. Magurran, and C. M. Garcia. 2005. Patterns of speciation in endemic Mexican Goodeid fish: sexual conflict or early radiation? *Journal of Evolutionary Biology* 18:922-929.

Robinson, S. K., and J. Terborgh. 1995. Interspecific Aggression and Habitat Selection by Amazonian Birds. *Journal of Animal Ecology* 64:1-11.

Roderick, G. K., and R. G. Gillespie. 1998. Speciation and phylogeography of Hawaiian terrestrial arthropods. *Molecular Ecology* 7:519-531.

Roelants, K., D. J. Gower, M. Wilkinson, S. P. Loader, S. D. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007. Global patterns of diversification in the history of modern amphibians. *Proceedings of the National Academy of Sciences of The United States of America* 104:887-892.

Rosenzweig, M. L. 1975. On continental steady states of species diversity. Pp. 121-140 in M. L. Cody, and J. M. Diamond, eds. *Ecology and Evolution of Communities*. Belknap, Cambridge.

Rubenstein, D. R., and I. J. Lovette. 2007. Temporal environmental variability drives the evolution of cooperative breeding in birds. *Current Biology* 17:1414-1419.

Ruber, L., J. L. Van Tassell, and R. Zardoya. 2003. Rapid speciation and ecological divergence in the American seven-spined gobies (Gobiidae, Gobiosomatini) inferred from a molecular phylogeny. *Evolution* 57:1584-1598.

Ruber, L., and R. Zardoya. 2005. Rapid cladogenesis in marine fishes revisited. *Evolution* 59:1119-1127.

Sanderson, M. J. 2002. Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology and Evolution* 19:101-109.

Sanderson, M. J., and M. J. Donoghue. 1996. Reconstructing shifts in diversification rates on phylogenetic trees. *Trends in Ecology & Evolution* 11:15-20.

Sanderson, M. J., and M. F. Wojciechowski. 1996. Diversification rates in a temperate legume clade: Are there "so many species" of *Astragalus* (Fabaceae)? *American Journal of Botany* 83:1488-1502.

Sargent, R. D. 2004. Floral symmetry affects speciation rates in angiosperms. *Proceedings of the Royal Society of London Series B-Biological Sciences* 271:603-608.

Sargent, R. D., M. A. Mandegar, and S. P. Otto. 2006. A model of the evolution of dichogamy incorporating sex-ratio selection, anther-stigma interference, and inbreeding depression. *Evolution* 60:934-944.

Sargent, R. D., and S. P. Otto. 2004. A phylogenetic analysis of pollination mode and the evolution of dichogamy in angiosperms. *Evolutionary Ecology Research* 6:1183-1199.

Sargent, R. D., and S. P. Otto. 2006. The role of local species abundance in the evolution of pollinator attraction in flowering plants. *American Naturalist* 167:67-80.

Savolainen, V., S. B. Heard, M. P. Powell, T. J. Davies, and A. O. Mooers. 2002. Is cladogenesis heritable? *Systematic Biology* 51:835-843.

Schluter, D. 1996. Ecological causes of adaptive radiation. *American Naturalist* 148:S40-S64.

Schluter, D. 2000. *Ecology of Adaptive Radiation*. Oxford Univ. Press, Oxford.

Seddon, N., R. M. Merrill, and J. A. Tobias. 2008. Sexually selected traits predict patterns of species richness in a diverse clade of suboscine birds. *American Naturalist* 171:620-631.

Seehausen, O. 2006. African cichlid fish: a model system in adaptive radiation research. *Proceedings of the Royal Society B-Biological Sciences* 273:1987-1998.

Sepkoski, J. J. 1976. Species diversity in the Phanerozoic: species-area effects. *Paleobiology* 2:298-303.

Sepkoski, J. J. 1978. A kinetic model of Phanerozoic taxonomic diversity I. Analysis of marine orders. *Paleobiology* 4:223-251.

Sepkoski, J. J. 1998. Rates of speciation in the fossil record. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 353:315-326.

Shaw, A. J., C. J. Cox, B. Goffinet, W. R. Buck, and S. B. Boles. 2003. Phylogenetic evidence of a rapid radiation of pleurocarpous mosses (Bryophyta). *Evolution* 57:2226-2241.

Sibley, C. G., and J. E. Ahlquist. 1990. *Phylogeny and classification of the birds of the world*. Yale University Press New Haven, CT.

Simpson, G. G. 1953. *The Major Features of Evolution*. Columbia University Press, New York.

Sims, H. J., and K. J. McConoway. 2003. Nonstochastic variation of species-level diversification rates within angiosperms. *Evolution* 57:460-479.

Skinner, A. 2007. Phylogenetic relationships and rate of early diversification of Australian *Sphenomorphus* group scincids (Scincoidea, Squamata). *Biological Journal Of The Linnean Society* 92:347-366.

Slowinski, J. B., and C. G. Guyer. 1989. Testing the stochasticity of patterns of organismal diversity: an imporoved null model. *American Naturalist* 134:907-921.

Stanley, S. 1975. A theory of evolution above the species level. *Proc. Natl. Acad. Sci. U.S.A.* 72:646-650.

Stanley, S. M. 1979. *Macroevolution: Pattern and Process*. Freeman, San Francisco.

Stanley, S. M. 2007. An analysis of the history of marine animal diversity. *Paleobiology* 33:1-55.

Stanley, S. M., K. L. Wetmore, and J. P. Kennett. 1988. Macroevolutionary differences between the 2 major clades of Neogene planktonic Foraminifera. *Paleobiology* 14:235-249.

Stephens, P. A., S. W. Buskirk, G. D. Hayward, and C. Martinez del Rio. 2005. Information theory and hypothesis testing: a call for pluralism. *Journal of Applied Ecology* 42:4-12.

Streelman, J. T., and P. D. Danley. 2003. The stages of vertebrate evolutionary radiation. *Trends In Ecology & Evolution* 18:126-131.

Stuart-Fox, D., and I. P. F. Owens. 2003. Species richness in agamid lizards: chance, body size, sexual selection or ecology? *Journal of Evolutionary Biology* 16:659-669.

Swofford, D. L. 2000. PAUP*. Sinauer, Sunderland, MA.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* 25:4876-4882.

Thorne, J. L., H. Kishino, and I. S. Painter. 1998. Estimating the rate of evolution of the rate of molecular evolution. *Molecular Biology and Evolution* 15:1647-1657.

Turgeon, J., R. Stoks, R. A. Thum, J. M. Brown, and M. A. McPeek. 2005. Simultaneous Quaternary radiations of three damselfly clades across the Holarctic. *American Naturalist* 165:E78-E107.

Valentine, J. W. 1969. Patterns of taxonomic and ecological structure of the shelf benthos during Phanerozoic time. *Paleontology* 12:684-709.

Valentine, J. W. 1985. Biotic diversity and clade diversity. Pp. 419-424 in J. W. Valentine, ed. *Phanerozoic Diversity Patterns*. Princeton Univ. Press, Princeton, NJ.

Valentine, J. W., D. H. Erwin, and D. Jablonski. 1996. Developmental evolution of metazoan bodyplans: the fossil evidence. *Developmental Biology* 173:373-381.

Valentine, J. W., and E. M. Moores. 1972. Global tectonics and the fossil record. *J. Geol*. 80:167-184.

Vamosi, J. C., and S. P. Otto. 2002. When looks can kill: the evolution of sexually dimorphic floral display and the extinction of dioecious plants. *Proceedings of the Royal Society of London Series B-Biological Sciences* 269:1187-1194.

Vamosi, J. C., S. P. Otto, and S. C. H. Barrett. 2003. Phylogenetic analysis of the ecological correlates of dioecy in angiosperms. *Journal of Evolutionary Biology* 16:1006-1018.

Vamosi, J. C., and S. M. Vamosi. 2004. The role of diversification in causing the correlates of dioecy. *Evolution* 58:723-731.

Vamosi, J. C., and S. M. Vamosi. 2005. Present day risk of extinction may exacerbate the lower species richness of dioecious clades. *Diversity and Distributions* 11:25-32.

Vamosi, J. C., and J. R. U. Wilson. 2008. Nonrandom extinction leads to elevated loss of angiosperm evolutionary history. *Ecology Letters* 11:1047-1053.

von Hagen, K. B., and J. W. Kadereit. 2003. The diversification of Halenia (Gentianaceae): ecological opportunity versus key innovation. *Evolution* 57:2507-2518.

Vrba, E. S. 1984. What is species selection? Systematic Zoology 33:318-328.

Walker, T. D., and J. W. Valentine. 1984. Equilibrium models of evolutionary species diversity and the number of empty niches. *American Naturalist* 124:887-899.

Wall, D. P. 2005. Origin and rapid diversification of a tropical moss. *Evolution* 59:1413-1424.

Webster, A. J., R. J. H. Payne, and M. Pagel. 2003. Molecular phylogenies link rates of evolution and speciation. *Science* 301:478-478.

Weir, J. T. 2006. Divergent timing and patterns of species accumulation in lowland and highland neotropical birds. *Evolution* 60:842-855.

Weir, J. T., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. *Proc. R. Soc. Lond. Biol. B*. 271:1881-1887.

Werdelin, L., and M. E. Lewis. 2005. Plio-Pleistocene Carnivora of eastern Africa: species richness and turnover patterns. *Zoological Journal Of The Linnean Society* 144:121-144.

Whittall, J. B., and S. A. Hodges. 2007. Pollinator shifts drive increasingly long nectar spurs in columbine flowers. *Nature* 447:706-712.

Wiens, J. J., and M. J. Donoghue. 2004. Historical ecology, biogeography, and species richness. *Trends Ecol. Evol.* 19:639-644.

Wiens, J. J., C. H. Graham, D. S. Moen, S. A. Smith, and T. W. Reeder. 2006. Evolutionary and ecological causes of the latitudinal diversity gradient in hylid frogs: Treefrog trees unearth the roots of high tropical diversity. *American Naturalist* 168:579-596.

Wiens, J. J., G. Parra-Olea, M. Garcia-Paris, and D. B. Wake. 2007. Phylogenetic history underlies elevational biodiversity patterns in tropical salamanders. *Proceedings Of The Royal Society B-Biological Sciences* 274:919-928.

Wiens, J. J., and T. A. Penkrot. 2002. Delimiting species using DNA and morphological variation and discordant species limits in spiny lizards (Sceloporus). *Systematic Biology* 51:69-91.

Williams, O. B., ed. 1982. *The vegetation of present-day arid Australia: a biotic appraisal*. Peacock Publications, Adelaide.

Williams, S. T., and D. G. Reid. 2004. Speciation and diversity on tropical rocky shores: A global phylogeny of snails of the genus Echinolittorina. *Evolution* 58:2227-2251.

Wollenberg, K., J. Arnold, and J. C. Avise. 1996. Recognizing the forest for the trees: testing temporal patterns of cladogenesis using a null model of stochastic diversification. *Molecular Biology and Evolution* 13:833-849.

Yang, A. S. 2001. Modularity, evolvability, and adaptive radiations: a comparison of the hemi- and holometabolous insects. *Evolution & Development* 3:59-72.

Yule, G. U. 1924. A mathematical theory of evolution based on the conclusions of J. C. Willis, F. R. S. *Proceedings of the Royal Society of London, Series B: Biological Sciences* 213:21-87.

Zamudio, K. R., and H. W. Greene. 1997. Phylogeography of the bushmaster (Lachesis muta : Viperidae): implications for neotropical biogeography, systematics, and conservation. *Biological Journal of the Linnean Society* 62:421-442.

Zhang, J. 1999. Performance of likelihood ratio tests of evolutionary hypotheses under inadequate substitution models. *Molecular Biology and Evolution* 16:868-875.

Zhang, L. B., H. P. Comes, and J. W. Kadereit. 2004. The temporal course of quaternary diversification in the European high mountain endemic Primula sect. Auricula (Primulaceae). *International Journal Of Plant Sciences* 165:191-207.

Zink, R. M., J. Klicka, and B. R. Barber. 2004. The tempo of avian diversification during the quaternary. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 359:215-219.

Zink, R. M., and J. B. Slowinski. 1995. Evidence from molecular systematics for decreased avian diversification in the Pleistocene Epoch. *Proceedings of the National Academy of Sciences of the U.S.A*. 92:5832-5835.